

When steady state is no longer good enough

The US Department of Energy will have to build something soon if it is to retain its role as provider of state-of-the-art scientific facilities. It should draw up a realistic medium-term plan for facility construction.

The dog's breakfast that makes up US science policy does have a certain rationale. The National Institutes of Health (NIH) and the National Science Foundation (NSF) provide grants to university scientists. The Pentagon supports computer science and engineering. The National Aeronautics and Space Administration holds televised press conferences. And the Department of Energy provides premium facilities, that no university — or even group of universities — could ever afford, to keep US science at the cutting edge.

Although these machines used to be primarily for physicists, their reach is increasingly broad. Neutron and splendid new light sources are useful for research in all kinds of materials science, and also make possible the study of organic molecules. Neutrons are especially powerful in this regard, as they interact well with light atomic nuclei.

The energy department has recently been confronted by opponents who would like to shut it down. Forced onto the defensive, it has balked at the cost of new machines. Although a stream of these are now coming on line, all were planned by the Bush administration. The Clinton administration stopped design work on a \$3-billion Advanced Neutron Source, the last and second most expensive of the facilities planned by Bush.

It is good that the department has succeeded in keeping the wolves from the door, and managed to defend its \$2-billion annual budget for civilian science. It has even consolidated that budget, with the administration acknowledging that it should be treated as well as NSF and NIH. But the Department of Energy's facilities are expensive to run, as well

as to build, and unless they remain state-of-art, these operating costs will become impossible to justify.

In striking a balance between construction, operations and actual science, the department has rightly taken steps to ensure that existing facilities are well used. Ernie Moniz, a physicist from the Massachusetts Institute of Technology then at the White House, helped to implement a Scientific Facilities Initiative which provided badly needed resources for operations. In a fiscally restrictive climate, the need to make better use of existing facilities was something on which the administration and the Congress could agree.

Now Moniz is returning to Washington, as under-secretary at the energy department (see page 510). The fiscal climate has improved somewhat, although not to the point where new construction programmes will be embraced without reservation.

But when countries such as Taiwan, Australia and Canada are planning neutron sources better than anything extant in the United States, it is time for the US Department of Energy to think about the long term. The neutron situation is particularly acute (see page 503), and should at least be addressed by funding for a new National Spallation Neutron Source in next year's budget proposal.

Beyond that, the department needs to get over the trauma that it has been through during the past few years, and lay out a realistic multi-year programme for the new facilities at its non-defence laboratories that will justify their existence into the next century. □

Double trouble

Unesco should think again before endorsing an outright condemnation of human cloning.

There has at times been something 'Kafkaesque' about the recent political debate on the ethics of human cloning. Politicians have been only too ready to make populist condemnations of a practice that still remains many years from realization. Others have found it convenient to leap on this moral bandwagon as a vehicle for delivering coded messages about an essentially separate issue, namely the morality of human abortion.

The latest victim of what appears at times to be virtually a witch-hunt has been the laudable efforts of the United Nations Educational, Scientific and Cultural Organization (Unesco) to draft a universal 'declaration on the human genome' (see page 508). Four years of debate have produced a thoughtful set of guidelines on protecting human rights in the genetic age. Last week, the draft text was modified — apparently under political pressure from some member states — explicitly to identify "reproductive cloning of human beings" as a practice "contrary to human dignity".

Much of the motivation behind this move is unexceptionable. If human cloning did become widely practised, it would undermine our thinking about the moral status of individuals, in particular our confidence in our 'uniqueness'. To challenge that uniqueness would have profound implications, especially if it was done for commercial

or political reasons. But is this sufficient reason to seek a blanket condemnation of the techniques themselves? Some would argue yes, on the grounds that anything less implies approval, opening a loophole that would soon grow much bigger.

There are several problems with this argument. One is that it allows no exceptions — including practices that could well meet Unesco's criterion of maintaining human dignity; an example is the possibility that it could enable a man with total germ-cell failure nevertheless to become a biological father.

A second objection is that outright condemnation, without a clear and unambiguous moral justification, may make good religion, but never makes good politics. Attempting to impose moral standards from above that lack the legitimacy of acceptance from below is always dangerous and misguided.

Those ultimately responsible for the final text should think again. After all, the guidelines will not themselves be legally binding, and the more rigidly they are written, the greater the likelihood that they may be ignored. To reject a blanket condemnation of human cloning, while insisting on a clear set of moral principles by which techniques such as cloning should be judged by individual countries and communities, would be an act of political courage, not cowardice. □



Delays at Brookhaven reactor worsen US 'neutron drought'

[WASHINGTON] The High Flux Beam Reactor (HFBR) at the Brookhaven National Laboratory in New York State, which was forced to close down in January after a radiation leak, is unlikely to reopen until October 1999 at the earliest, according to the US Department of Energy (DOE).

The prolonged mothballing of the reactor will further exacerbate an existing shortage of neutron beam lines in the United States, say scientists, crowding existing facilities and reducing the scope for innovative experiments that need plenty of beam time.

In the longer term, there is concern that the lack of facilities will drive researchers from the field and lead to a growing divergence between Europe, which is thought to have around 3,000 regular neutron users, and the United States, with perhaps 1,000.

Instead, US scientists who want to study the structure of matter — including chemists, condensed-matter physicists and structural biologists — will have to rely on photons, which are more readily available.

Neutron scientists, who tend to believe that the advantages of neutron scattering are self-evident, are now engaged in a struggle to persuade the DOE and the Congress that this outcome will have an adverse impact on US science as a whole.

But when the Basic Energy Sciences Advisory Committee (BESAC), a panel of external scientists that advises the energy department, met near Washington last week to begin preparing a report on the need to restart HFBR (see box, next page), it was apparent that some members, at least, have still to be convinced that this is so.

The panel, which is chaired by John Stringer of the Electric Power Research Institute (EPRI), Palo Alto, California, has been charged with identifying the impact of the HFBR shutdown, as well as advising on

whether it should be restarted. It is almost certain to call for the reopening of the reactor when it delivers its report in October.

But such a call will not ensure the future of the reactor. And even if it does restart, neutron science in the United States will remain in an uncertain state. Indeed, since 1994 when the US discontinued work on the Advanced Neutron Source (ANS) — a \$3-billion reactor-based neutron source planned for the Oak Ridge National Laboratory in Tennessee — that condition has steadily worsened.

As Europe proceeds with plans for several upgraded and new facilities to address what has been described as a 'neutron drought' (see *Nature* 379, 284; 1996) — including the FRM-II reactor under construction at Munich, and the planned \$1-billion European Spallation Source — and aspirant scientific powers as far apart as Taiwan and Austria plan substantial investment in their own neutron sources, the United States is struggling just to keep what it has.

Even before a leak of tritium closed Brookhaven's HFBR (see *Nature* 386, 3; 1997), the energy department had rejected proposals for the extensive refurbishment of both that reactor and the High-Flux Isotope Reactor (HFIR) at Oak Ridge, Tennessee.

Instead, the department has embarked on improvements to instrumentation at HFIR, and an upgrade of the neutron spallation source at the Los Alamos Neutron Science Center (LANSCE) in New Mexico, which has been able to cut costs by using a particle accelerator paid for by nuclear weapons research dollars.

The department is still pursuing the design of a much larger spallation source, the National Spallation Neutron Source (NSNS), at Oak Ridge. But the \$1.3 billion it would cost to build the NSNS is not yet

assured, and it will produce no neutrons until the year 2005 at the earliest.

For the 270 former users of the Brookhaven reactor, the DOE has good and bad news. The good news is that there is a plan for restart, and outright opposition to the reactor from the local community may be on the wane. The bad news is that this plan will not get the reactor going until October 1999, and that potential pitfalls could push this date back into the next century.

The plan would see the reactor restart at its previous operating power of 30 MW and later upgrade it to 60 MW, the full power at which it operated for a period in the 1980s. Brookhaven scientists are keen to operate at the higher power, as it gives them twice as many neutrons for about 10 per cent extra operating cost.

But the plan would require an Environmental Impact Statement (EIS), to assure the public that a restart decision was safe, before making the structural improvements needed to bring the facility into line with current earthquake-proofing standards.

The department is split over the need for the EIS, which it says it can carry out in 15 months. Some officials point out that the complex, \$2-million study is unnecessary to restart the reactor legally. But the prevailing view is that the EIS is needed to reassure the public and — perhaps more crucially — to stave off legal challenges from environmental groups, which might otherwise demand a last-minute EIS, delaying the project further.

Federico Peña, the Energy Secretary, has said he will personally decide in January of next year if the department should pursue the restart option, on the basis of advice from the Long Island community, from the new contractor at Brookhaven — which is expected to be appointed in November — and from BESAC.

The EIS and then the construction work have to be finished before neutron science can resume at Brookhaven. Pat Dehmer, head of Basic Energy Sciences at the energy department, says she is confident the schedule will hold, citing Peña's personal commitment to it.

Sceptics point out, however, that the new contractor may balk at making a recommendation on the safety of the reactor within two months, as the schedule requires. They also point out that Peña is widely expected to be replaced by his deputy and heir-apparent, Elizabeth Moler, by the end of 1998; that the department has never previously completed an EIS in 15 months; and that the EIS would itself be open to legal challenge.



No neutrons today: closure of HFBR at Brookhaven is forcing researchers to look for other sources.

Nonetheless, prospects for the reactor have improved since Brookhaven's crisis peaked in March, at which time some scientists feared the reactor would never reopen. "It is now clear that there are numerous voices, across the community, for restart," says Martha Krebs, assistant secretary for energy.

But BESAC's debate over restart demonstrates how far the expectations of the US neutron community have fallen. According to Lyle Schwarz, chair of Associated Universities Inc., the sacked Brookhaven contractor, the community has always said it needs the new NSNS as well as the existing reactor sources. "I find it strange that we're even discussing the question [of whether HFBR is needed]," he says.

Prospects for the construction of the NSNS have improved somewhat this summer. Both houses of Congress have allowed the administration's request of \$23 million next year for design work on the project. After a detailed review by 60 outside scientists who visited Oak Ridge last month, Dehmer says that money to start construction will be in the 1999 budget, which President Bill Clinton will propose next February.

But the review estimated the cost of NSNS at \$25 million higher than the \$1.265 billion proposed by the design team, and even the latter figure has crept up from the \$1 billion which some observers had expected the facility to cost. The review also ques-

Table 1 **Comparison of main US and European reactor neutron sources**

Reactor	Power	Flux (10^{14} ntrns/cm ² s)	Instruments	No. of users	Operating cost (annual)
United States					
High-Flux Isotope Reactor (HFIR)	85 MW	5	13	200	\$29m.
High-Flux Beam Reactor (HFBR)	30 MW	10	17	270	\$26m.
NIST Centre for Neutron Research	20 MW	4	25	750	\$6m.
University of Missouri Research Reactor	10 MW	1	7	[60]	\$8m.
Europe					
Institut Laue Langevin (ILL)	57 MW	12	32	1,500	\$26m.
FRM-II (under construction)	20 MW	8	25	[-]	\$25m.

tioned whether it could be built in six years. Any stretch in that timescale would add further to its cost.

NSNS benefits from its proposed location in the home state of Vice-President Al Gore, who is also the favourite to serve as president during its construction phase. Objections to the facility's location from Gore's political opponents in the Congress have been assuaged by a strongly collaborative management plan that will share the project with laboratories outside Tennessee.

In the short term, however, the outlook for neutron availability is bleak. Upgrades will create a better spallation source at Los Alamos and a more flexible reactor at Oak Ridge by the year 2000. But in each case, the modifications will curtail research activity between now and then.

The most popular neutron facility in the United States, at the National Institute of

Standards and Technology (NIST) at Gaithersburg in Maryland, has made some of its capacity available this year to researchers displaced from Brookhaven. The 20-MW NIST reactor has many instruments and a large and diverse user base, but it provides a lower neutron flux than either Brookhaven or Oak Ridge. "I have high-quality science going on which I am displacing to help meet a short term-need," says Mike Rowe, director of the NIST reactor. "I can't keep doing it forever."

The University of Missouri at Columbia, meanwhile, has stepped forward with a proposal to alleviate the shortage that some laboratory officials did not want to hear. According to Missouri's Bill Yelon, a 10-MW reactor at the university could take one-fifth of Brookhaven's users, if the department gives it \$500,000 for instruments and another \$500,000 a year in running expenses. Yelon hopes that this support would help to persuade his university to build a \$28-million facility that is needed for housing extra neutron experiments.

Iran Thomas, deputy director of basic energy sciences at DOE, says that backing for Missouri would bring pleas for support from all the other US university reactors. But several BESAC members like Yelon's proposal. "What's wrong with that as an interim solution?" enquired panel member Jan Herbst, of General Motors.

The answer, unfortunately, is that politics will preclude the DOE from shopping around, either in the United States or abroad, to fill its neutron gap. The cancellation of the ANS, the leak at Brookhaven, and the still precarious status of the proposed spallation source add up to a *de facto* deprioritization of neutron science by the energy department.

While investigators crowd on to the United States' lavish new photon-based facilities, the neutron community will have to fight for space on the facilities remaining. BESAC does not seem impressed by the argument that this is unacceptable: it sent the community leaders off to a room to identify the weak science that will be lost under existing capacity constraints. The EPRI's Stringer hopes the outcome will help BESAC make a compelling case for the restart of Brookhaven. "The decision we're looking for is a positive one," he says. **Colin Macilwain**

Women researchers take on 'old boys' network

[WASHINGTON] Women members of the US Department of Energy's Basic Energy Sciences Advisory Committee (BESAC) have challenged what they see as the remnants of an 'old boys' network' pattern of neutron use that risks circumventing the peer review process.

At last week's meeting of BESAC (see above) several women questioned whether DOE facilities were really open to all scientists on the basis of merit, or whether people with the right connections monopolize beam time at the expense of outsiders. And the women suggested that their gender is heavily underrepresented among neutron users.

Patricia Dehmer, director of the department's \$700-million basic energy sciences programme, has transformed the composition of the advisory panel.

In doing so, she has not

only shifted the gender balance but also demolished the cosy consensus that used to exist between the panel and the programme managers it is supposed to supervise.

Last week's collision arose when Jack Fischer, vice-chair of the Neutron Scattering Society of America, told the panel how researchers with contacts at the facilities can often find ways to explore ideas that they consider interesting.

Geraldine Richmond, a chemist at the University of Oregon and two-year member of BESAC, asked why this was going on, when "this community doesn't have a reputation of generosity to outside users". Fischer responded that good ideas "didn't go away" just because they failed the peer review process.

At that point, Marye Anne Fox, vice-president for

research at the University of Texas at Austin, and the most senior female scientist on the panel, rebuked Fischer for failing to treat Richmond's question seriously and accused him of condoning "a circumvention of the peer review process".

Patricia Thiel, head of materials chemistry at Iowa State University, then asked Fischer for a breakdown of the Neutron Scattering Society's membership. An office-bearer of the society said that women members numbered 15-20 out of a total of 750.

As Thiel observed, this suggested women are grossly underrepresented among neutron users, as compared with their representation in fields of science such as structural biology and many branches of chemistry that could use the technique. **C.M.**

Backing for anti-cloning bill reopens embryo debate

[WASHINGTON] The science committee of the US House of Representatives sent a shot across the bows of the biomedical research community last week by endorsing an anti-cloning bill significantly rewritten to ban the cloning of human embryos — not merely the cloning of ‘human beings’.

The decision was the first action on cloning legislation by a congressional committee. It immediately drew complaints that it ignored the recommendations of President Bill Clinton’s National Bioethics Advisory Commission, needlessly reopened the debate on human embryo research, and risked restricting beneficial research that used cloning technology but was not aimed at producing cloned humans.

“The US seems hell bent on withdrawal from one of the most important areas of biomedical research,” wrote David Blake, vice-president for biomedical research at the Association of American Medical Colleges, in a message sent on the Internet to advocates of biomedical research. Sean Tipton of the American Society for Reproductive Medicine, agreed: “We are very concerned that this bill, rather than a cloning bill, has become a permanent embryo research ban.”

“The debate has now totally changed,” says Chuck Ludlam, vice-president for government relations at the Biotechnology Industry Organization (BIO). In a letter to Lynn Rivers (Democrat, Michigan), a Congress science committee member, Ludlam said he feared that Vernon Ehlers (Republican, Michigan) would use the bill as “a vehicle for relitigating the embryo research issue”. This is a reference to a perennial debate in the Congress between opponents of abortion, mainly Republican, who seek to outlaw human embryo research, and research advocates, mainly Democrats, who say that such research should be permissible within careful constraints.

Ehlers, a physicist by training, argues that his bill will in fact protect biomedical research

by pre-empting more drastic bills which could lead to a ban on cloning in all research, including animal work. Ehlers told *Nature* that the rewriting was dictated by political necessity, as a “majority” of science committee members, led by chairman James Sensenbrenner (Republican, Wisconsin), “said they couldn’t vote for the bill without also banning cloning of embryos”.

The rewritten bill bans federal funding for “any project of research that includes the use of human somatic cell nuclear transfer technology to produce an embryo”. When Ehlers first introduced the bill in March, it banned funding for work using “a human somatic cell for the process of producing a human clone”.

Most political sympathies in the committee were with the rewritten bill. When Rivers proposed an amendment to restore the ban to one on the use of cloning for “creation of a human being”, it was soundly defeated. Another Rivers amendment did pass, protecting the use of cloning technologies to clone molecules, DNA, cells, tissues or cells other than human embryo cells, and to create animals other than humans.

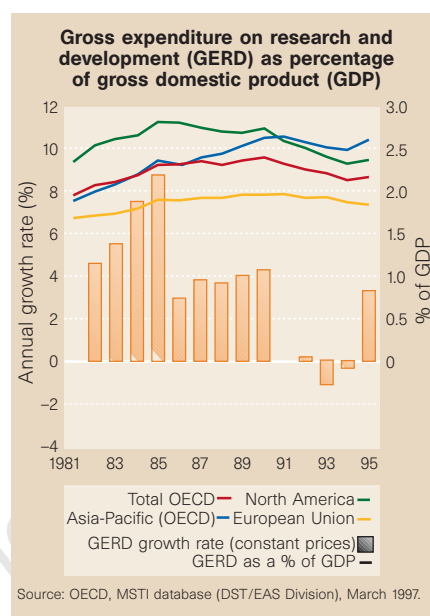
But BIO and others still argue that the bill as passed could jeopardize broad classes of research such as the growth of lines of nerve, liver or kidney cells. The reason, according to BIO, is that in the ‘post-Dolly’ era, any somatic cell can in theory be “used ... to produce” an embryo. So a scientist working with a cell that could potentially give rise to an embryo could be open to having his intent questioned and possibly fall within the bill’s prohibitions.

In addition, the fact that the law would be permanent would create “the worst of all worlds”, said Ludlam. President Clinton and the bioethics advisory commission recently supported anti-cloning legislation that does not outlaw the use of embryos to create research embryos if they are not implanted, but that bill would expire after five years.

But what BIO fears most is a second Ehlers bill that applies not to federally funded scientists but to the private sector. That bill, now before the House of Representatives commerce committee, seeks a \$5,000 penalty for anyone, public or private, attempting to produce a ‘human clone’. Ehlers has not rewritten the bill to penalize the cloning of a human embryo, but may yet make such a change.

Federal funding for human embryo research is already banned under provisions attached to the spending bills that fund the National Institutes of Health. The Ehlers bill would make the ban permanent for the use of cloning technologies in creating embryos, though not for embryo research not involving cloning.

Meredith Wadman



Japan leads rise in research budgets in OECD countries

[PARIS] Research budgets measured collectively across the world’s industrialized countries are increasing again after the economic slump of the early 1990s, according to latest figures from the Organization for Economic Cooperation and Development (OECD).

Between 1994 and 1995, the total spending of the OECD’s 27 member states on research and development increased from 2.1 to 2.2 per cent of their joint gross domestic product (GDP), the first increase since 1990, when it peaked at 2.4 per cent (see chart).

The OECD statistics show that the most research-intensive member states are clearly those in the Asia-Pacific region, particularly Japan which, despite its economic difficulties, spent 2.8 per cent of its GDP on research and development in 1995.

But there has also been a slight upturn in research spending in North America, led by the United States, after an almost continuous decline since the mid-1980s; the OECD warns, however, that forecast data for 1996 “did not suggest continued growth”. The European Union’s record has been mixed, with overall spending dropping slightly — from 1.9 to 1.8 per cent of GDP — between 1994 and 1995.

Among those whose research and development profile the OECD notes has changed significantly, Ireland spent twice the percentage of GDP on research and development in 1995 as in 1981, and recorded high growth rates in expenditure, ranging from 15 to 20 per cent a year, in the 1990s.

Science, Technology and Industry: Scoreboard of Indicators 1997. OECD, Paris. FF200. ISBN: 92-64-15507-4



US grants support teraflop computing simulations

[WASHINGTON] The US Department of Energy (DOE) last week announced grants worth \$250 million over 10 years to five universities to help to develop the computational aspects of its effort to simulate phenomena associated with nuclear weapons.

Federico Peña, the Secretary of Energy, announced the selection of California Institute of Technology, Stanford University, University of Chicago, University of Illinois at Urbana/Champaign and University of Utah as 'centres of excellence' under the department's Academic Strategic Alliances Program (ASAP).

The programme is part of the larger Accelerated Strategic Computing Initiative, the department's plan to push the state of computational science beyond 100 teraflops by 2006 (a teraflop is a million million floating point operations per second).

The grant awards, selected from 49 proposals, will support the development of algorithms and computational approaches by the universities to help them to solve related, but unclassified, problems. The researchers will have access to three teraflop-scale computers at the department's nuclear weapons laboratories. About 10 per cent of the computers' operating time is expected to be devoted to university work, and converting the computers from classified to unclassified operation can take as little as 20 minutes.

Faculty and postdoctoral graduate researchers at each of the universities will receive about \$5 million annually to conduct research on physics-based modelling and high-performance computer simulation.

The problems to be addressed by the universities appear as relevant to industrial applications as to nuclear weapons research. Stanford, for example, will study the complex turbulent flows in aircraft jet engines.

Utah will work to provide a set of science-based tools for numerical simulation of accidental fires and explosions. And the California institute will investigate the effect of shock waves induced by high explosives on various materials in different phases, work that will be useful in areas such as mine accident rescue and building demolition.

Perhaps the most exotic winning proposal was that from Chicago. According to David Schramm, the university's vice-president, it will seek to unravel the astrophysical mystery of how a supernova explodes. Schramm says that, unlike the DOE, he will be able to continue testing his simulations in space.

John Hennessy, dean of engineering at Stanford, says the DOE computers are roughly two orders of magnitude more powerful than any now available to the university.

Victor Reis, assistant secretary of energy for defence programmes, says the DOE did not specify in the competition the research areas to be addressed. He says the aim is simply to "push" high performance computing, and make it work more efficiently.

C. Paul Robinson, director of Sandia National Laboratories, home to the first teraflops machine, says the main objective of ASAP is to develop algorithms and "equation solvers" to enable high-powered computers to run simulations as efficiently as possible.

David Kramer

Modest increase proposed for EU research budget

[MUNICH] The European Commission last week proposed a small increase in the funding of research in the European Union's fifth multi-year Framework programme (FP5), due to start in 1999, to a total of ECU16.3 billion (US\$17.5 billion).

The commission's draft proposal for FP5, published in April, suggested concentrating research within three thematic programmes, each absorbing 24 per cent of an unspecified total budget. The remaining money would be divided between the union's nuclear research programme and three policy-orientated programmes, to support, for example, mobility of researchers.

The proposed budget represents an increase of about 3 per cent, over that of the current fourth Framework programme, of its proportion of the total gross national product of member states. But this may not be enough to satisfy the European Parliament, which must approve the programme and its budget jointly with the Council of Ministers.

Several years ago, the parliament demanded that the percentage of the commission's budget devoted to research should be increased from 4 to 6 per cent. It is not yet clear whether the commission's budget proposal represents a real increase in this factor.

Parliament's research committee presents its formal response to the FP5 proposal in mid-September. It is likely to request inclusion of an additional thematic programme on energy and the environment, and could also demand an even higher budget. The Council of Ministers is likely to be less enthusiastic about such a move.

Alison Abbott

Endangered species bill faces battle against property lobby

[WASHINGTON] A bill introduced in the US House of Representatives last week to reauthorize the Endangered Species Act (ESA) has been welcomed by environmentalists. But it is unlikely to settle deep differences with 'property rights' advocates.

The Endangered Species Recovery Act of 1997 (H.R. 2351), proposed by George Miller (Democrat, California), is the first ESA reauthorization bill to emerge in this session of Congress. The original act expired in 1992 after 20 years. Although its legal protection of endangered species remains, both environmentalists and property rights groups have been pressing for a new bill to remedy what they see as its shortcomings.

Miller's bill would make the primary goal of legislation the recovery of an endangered species to healthy status, rather than just the prevention of its extinction. That shift alone

is likely to make it unacceptable to those who complain that the current act already places excessive restrictions on landowners.

Equally divisive is whether private owners of protected habitat who agree to conservation plans should be required to alter, or even scrap, development plans if the needs of a species changes. The Clinton administration, led by the Interior Secretary, Bruce Babbitt, has promoted such Habitat Conservation Plans, more than 400 of which have been either approved or proposed.

Babbitt has further promised landowners that once agreement is reached on such a plan, the government could not impose further restrictions for as long as 100 years. But many scientists, and virtually all environmental groups, oppose this 'no surprises' policy, saying it is not flexible enough to accommodate new scientific

information (see *Nature* 386, 530; 1997).

Rather than strictly guaranteeing no surprises, Miller's bill calls for bonds, trust funds and tax incentives to help alleviate any financial burdens on landowners caused by compliance with the act.

On the same day that Miller introduced his legislation, Babbitt met with four Democratic and Republican senators who have been trying for months to negotiate a compromise on the new act.

Babbitt is said to have encouraged them to settle their differences on 'no surprises' and other issues, such as whether government agencies need to consult each other when determining the effects of land use on endangered species. But, says an aide to Senator Dirk Kempthorne (Republican, Idaho), a key figure in the negotiations, no solution is yet in sight.

Tony Reichhardt

New DFG head vows to back Germany's young scientists — and genetics research

[MUNICH] Concern for Germany's future scientists and a strong commitment to European collaboration in basic research are likely to characterize the three-year term of Ernst-Ludwig Winnacker (above) as president of the Deutsche Forschungsgemeinschaft (DFG) research council.

His appointment means that the DFG, which distributes grants to Germany's university researchers, will also remain firmly in egalitarian hands, as Winnacker — like his predecessor, Wolfgang Frühwald — believes that the entrenched traditional hierarchical culture of German science is counterproductive.

An internationally respected molecular virologist, Winnacker, who takes up his post at the beginning of next year, received his PhD from the Swiss Institute of Technology in Zurich in 1968. He was the first to clone the nuclear factor (NF1) gene involved in the control of transcription.

After pursuing research at the Karolinska Institute in Stockholm, at the University of California, Berkeley, and at the Institute of Genetics in Cologne, he was appointed to a chair in biochemistry at the Ludwig Maximilians University in Munich, where he has spent the past 20 years, and is currently director of the Genzentrum München.

Although Winnacker's first career choice was music — he plays the cello and the piano — he says he has no regrets about abandoning this for science. But life as a molecular biologist has not always been easy in Germany, where genetic engineering has faced aggressive and articulate opposition.

His energetic participation in public debates in defence of research, and a parliamentary advisory committee on genetic engineering from 1984 to 1987, resulted in Winnacker being placed on a hit-list by an extremist group, and he was under police protection for two years.

Despite the recommendations of the parliamentary committee, Germany introduced heavily restrictive rules on genetic engineering in 1990. But since then genetic engineering has been slowly gaining public acceptance, at least in medicine, and the rules were relaxed in 1993.

Winnacker describes the DFG as being "in excellent shape". But he has many ideas about areas of DFG policy where he would like to see improvements. At the top of the list is the possibility of expanding the council's programmes for young scientists.

He is keen to address the scope for awarding grants for periods of perhaps up to five years, in addition to the two-year norm. This would give research continuity, and help to reverse the trend towards



excessive reviewing, he says. "Sometimes I feel that we are spending all our time reviewing each other rather than working."

Winnacker also wants to see whether the large number of DFG specialist review committees — there are 37 — could be reduced or restructured to promote interdisciplinary approaches to science.

Winnacker has strong views on European research policy. He says that although the frequently proposed creation of a single European research council for basic research would be premature, given differences in research infrastructure between rich and poor countries, national research organizations should allocate some of their budgets to a joint fund to support collaborative projects (as the DFG already does with Israel and Palestine).

He also wants the DFG to become more directly involved in promoting public understanding of science. Part of the reason that genetics has had a bad press in Germany, he says, is that scientists do not describe to people what they actually do.

He practises what he preaches. He is the author of two popular books on genetics, and is working on a children's book about the life of a gene.

Winnacker considers his most important achievement to be the creation of the Genzentrum, one of three genetics centres established in the early 1980s with federal research ministry funding. The centre's building on the edge of Munich houses 80 scientists. The site is next to two Max Planck research institutes, and near Munich's largest teaching hospital. It will soon be joined by the university's entire faculty of science. This concentration of scientists was part of Winnacker's goal: "Critical mass is fundamental for successful science," he says, arguing that faculties in German universities tend to be too small.

Alison Abbott

Energy laboratories to open joint DNA sequence 'factory'

[SAN FRANCISCO] The US Department of Energy's Joint Genome Institute, a programme combining genomic research at three national laboratories, is to set up a high-throughput DNA-sequencing factory.

The factory, a joint effort by the Lawrence Livermore, Lawrence Berkeley and Los Alamos national laboratories, will be based in Walnut Creek, California, in the San Francisco Bay area. It represents a commitment to increase the three laboratories' screening capacity and so create a major contributor to the worldwide effort to sequence the human genome by the year 2005.

About 200 technicians and researchers will work three shifts around the clock, employing state-of-the-art robotics. Technology development and other research will continue at each laboratory. "The purpose is to [make sequencing] highly efficient and very cheap," says Elbert Branscomb, scientific director of the Joint Genome Institute.

The Department of Energy was the first federal agency to fund an initiative to unravel the entire genome. The impetus for this move grew out of its interest in detecting rare genetic changes that may result from exposure to toxic or radioactive substances. Until now, however, the national laboratories have lagged behind other major sequencing laboratories.

By the end of the current fiscal year in September, the three laboratories are likely to have contributed only about 3 million bases to the public databases, out of a total of 3 billion bases in the human genome. They hope, however, to increase that figure to 20 million by the end of next year, comparable to the most productive laboratories in the country.

The Joint Genome Institute, founded last year, represents a radical change in the laboratories' way of working, says Branscomb, with a shift from competition to collaboration between them. "The essence is breaking down the barriers between the labs," he says.

Historically, biological research programmes at the laboratories have been small compared with the applied physics and chemistry on which they built their reputations. The Joint Genome Institute has requested \$40 million for programmes next year at the three laboratories and the sequencing factory, a \$10 million increase on the current fiscal year.

Branscomb says that the sequencing factory will adhere to stringent public disclosure procedures on its work, with immediate postings on the web regarding smaller pieces, and full clones submitted nightly to the public databases.

Sally Lehrman

Unesco text will target gene techniques

[WASHINGTON] A meeting of government experts held at the United Nations Educational, Scientific and Cultural Organization (Unesco) in Paris has modified a proposed declaration on genetics and human rights, due to be debated this autumn, so that it explicitly denounces human cloning and germline gene therapy.

The modification goes beyond the advice of Unesco's International Bioethics Committee (IBC), which has spent four years drafting the original declaration, and whose 53 members had broadly urged that the declaration should stick to general principles and not address specific technologies.

Most of the draft declaration was endorsed by the meeting with only minor changes. But the modification on cloning and germline gene therapy has raised concern in some quarters that it would put a novel restriction on particular types of research and medical intervention.

"Putting particular technologies into a document of this kind is a mistake," says Harold Edgar, IBC member and professor of law, science and technology at Columbia University in New York. He says such declarations should "espouse relatively few principles at relatively high levels of generality".

But the changes may not have gone far

enough for some countries. Canada pulled out of last week's meeting, claiming in a statement that there had been "insufficient time to deal with unresolved articles and new substantive articles, such as human cloning and germline intervention". The Canadian delegation has come under pressure from aboriginal and disabled rights groups to take a strong stand on human rights.

The move to identify cloning and germline modifications as particular threats was led by delegates from Germany, and gained widespread support at the meeting, which was attended by experts from 83 countries.

"Anything that is not banned is allowed," declared Gerhard Fulda, Germany's ambassador to Unesco. "If we were to agree on a text which does not even mention a ban on cloning, we all know what the next day's headlines would be: 'Unesco: cloning still allowed,'" he told the meeting. "In Germany, we could not justify that politically."

The draft now states that "practices which are contrary to human dignity, such as reproductive cloning of human beings, shall not be permitted". It identifies germline intervention as being potentially among such practices, and urges the IBC to point this out to Unesco's general conference when recommending how to implement the accord.

But Germany failed to gather sufficient support for a further change prohibiting human embryo research. Germany lost a similar battle in the Council of Europe last November, when that body adopted a Convention on Human Rights and Biomedicine allowing the use of human embryos for research, provided they are not created for such purposes (see *Nature* 384, 298; 1996.).

A Unesco press release quotes Noëlle Lenoir, a member of the French Constitutional Council and the president of the IBC, as saying that the declaration "does not aim to regulate, authorise or restrict specific scientific processes", but rather "to establish lasting ethical principles at a universal level".

Nevertheless, Lenoir calls the new draft "appropriate", as it indicates that cloning "raises great difficulties" with respect to the principles of the declaration. She says the rewording does not change an essential balance in the text between the protection of human rights and scientific freedom.

David Shapiro, recently retired executive secretary of the UK Nuffield Council on Bioethics, and an IBC member, suggests that, although the inclusion of the cloning phrase could be seen as "banner waving", if it "is the price of getting agreement, I think that's well worth paying". **Meredith Wadman**

Australian policy review backs huge cuts in research funding

[SYDNEY] An attack on research funding in a government-commissioned report on reforming policy on industry has rocked Australia's scientific community. One proposal in particular, to slash government funding for the country's 65 Cooperative Research Centres (CRCs) by 70 per cent from July 1998, is described as "devastation" by the association that represents the centres.

John Moore, Minister for Industry, Science and Tourism in the Coalition government, commissioned the report from David Mortimer, chief executive of the transport company TNT, giving Mortimer the brief of reviewing the government's A\$4-billion (US\$3-billion) annual support for business. The one-man review comes at a time when public support for government has been falling sharply, fanned by criticism of the lack of a clear policy for industry.

Mortimer has proposed that government set a clear goal of doubling the rate of economic growth per capita over 10 years, and streamlining assistance schemes for industry into five key programmes, including innovation, competitiveness and sustainable management of resources.

The proposals have been largely welcomed by business groups. Within the government, however, stout resistance is



Mortimer: heavy cuts envisaged.

emerging from the Treasurer, Peter Costello — a sharp critic of continued subsidies for industry. Moore favours Mortimer's proposals, paving the way for a fierce ideological fight.

While attacking last year's unpopular cuts in tax concessions for industrial research and development from 150 to 125 per cent, Mortimer proposes that a further reduction in concession to 100 per cent should be counterbalanced by rebates to companies for expenses on innovation — an effective total of 137 per cent.

Government support for applied research would be reduced and the funding gap filled by industry, a suggestion of which many scientists are critical. "There is no argument or analysis to justify a wholesale dismantling of the public research effort," says Joe Baker, president of the Federation of Australian Scientific and Technological Societies.

The Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the CRC programme would be hit hardest by the recommended cut in support.

Mortimer has proposed that CSIRO's

external earnings target be increased from 30 to 50 per cent. But the chief executive of CSIRO, Malcolm McIntosh, says Mortimer used "the wrong base" for resetting this target. "His target means 70 per cent of science would have to come out of annual government appropriations," says McIntosh. "This would cripple strategic research and destroy CSIRO."

The CRC scheme is defined as an industry programme in the proposals, and Mortimer has suggested that the government's annual support of A\$146 million be cut to A\$20 million and limited to 'public good' centres. But Mark Sceats, deputy chair of the CRC Association, says the CRCs were shocked, as their scheme was outside Mortimer's brief. Peter McGauran, the Science and Technology Minister, says only that he has "strong views" on the CRC proposals that he plans to convey to the cabinet, of which he is not a member.

Mortimer also proposes that universities — already hit by government cuts — should boost income from business by 50 per cent. Sir Gustav Nossal, president of the Australian Academy of Science, says research cuts are a "by-product" of Mortimer's plans for industry. He has "failed to distinguish basic and strategic research from applications", says Nossal. **Peter Pockley**

EU official ordered to retire after claims of withholding data

[MUNICH] The European Commission's deputy director-general for agriculture with responsibility for veterinary affairs, Fernando Mansito Caballero, has been ordered to take early retirement. The order comes five months after accusations that he deliberately withheld from the public information about the dangers of bovine spongiform encephalopathy (BSE) in 1990.

The move is widely believed to result from pressure from the European Parliament, which had accused the commission of mismanaging the BSE affair and demanded an investigation within the commission. Mansito allegedly said in 1990 that BSE should be discussed behind closed doors to avoid inciting public alarm.

But the commission is denying that Mansito's forced resignation is related to these accusations. "The retirement of Señor Mansito took place within an internal reorganization. It is not a sanction," says a commission spokesman. "He will receive financial compensation and will be entitled to his full pension." Indeed, one official's report of a meeting in June 1990 records that Mansito reprimanded the members of the commission's scientific veterinary committee for playing down health risks and for "only discussing free circulation [of products] and economic aspects" (see *Nature* 385, 6; 1997).

Siberian hunters uncover mammoth remains

[MOSCOW] Scientists in Siberia were hurriedly called out last month to the banks of the Maksunuokha river in the Ust-Yansky region, where local hunters had discovered the frozen remains of a male mammoth. The scientists, from the biological resources department of the nature protection ministry in Yakutiya and from the local mammoth museum, found a carcass perfectly preserved in the permafrost, although it lacked tusks and part of its head.

The discovery has raised hopes that it may be possible to supply the mammoth sperm ordered several years ago by Japanese geneticists, hoping to cross-breed a Yakutian mammoth with a female African elephant.

Funds earmarked for AIDS vaccine centre

[WASHINGTON] Congress has specified that \$26.1 million of 1998 funding for the National Institutes of Health (NIH) be spent on a new AIDS vaccine research centre on the NIH campus in Bethesda, Maryland. The facility will house the collaborative vaccine

research efforts of the National Institute of Allergy and Infectious Diseases and the National Cancer Institute.

President Bill Clinton announced the new centre in May, when he called for the development of an AIDS vaccine within a decade. The money for the centre will come from NIH's budget for buildings and facilities, which has been allocated \$223 million in 1998 by the House of Representatives appropriations committee and \$212 million by its Senate equivalent.

UK group formed to help university contract staff

[LONDON] A group has been set up in Britain to improve the career management of university contract research staff. The main role of the Concordat Implementation Group will be to find practical ways to fulfill an agreement reached last year between universities and research-funding organizations on securing employment conditions for contract research staff that are in line with those for permanent employees (see *Nature* 383, 374; 1996).

The group will be chaired by Sir Gareth Roberts, vice-chancellor of the University of Sheffield and the current chairman of the Committee of Vice-Chancellors and Principals, one of the bodies responsible for the concordat. Other members will be drawn from key areas in higher education, as well as labour unions and representatives of research-funding bodies.

EU-Russian science agreement proposed

[MUNICH] The European Commission has proposed a formal agreement for scientific and technological cooperation with Russia, allowing Russian scientists to participate in the commission's Fifth Framework Programme (FP5), due to start in 1999. The agreement would allow Russian scientists to compete with scientists from European Union countries for participation in any FP5 project — outside the nuclear programme — but they would not receive funds.

The commission has formally requested authorization from the Council of Ministers to enter negotiations with Russia. The agreement would also clarify rules on intellectual property rights, and try to ensure that costs associated with Russian participation would be exempt from taxes.

MIT physics professor nominated for DOE

[WASHINGTON] Ernie Moniz has been nominated by President Bill Clinton to become under-secretary of energy, the number three position at the Department of Energy. Moniz is chair of the physics

department at Massachusetts Institute of Technology, and until recently was associate director for basic research at the US Office of Science and Technology Policy.

His nomination, predicted earlier this year (see *Nature* 387, 116; 1997), came after Federico Peña, the energy secretary, became convinced he should appoint a science adviser with real administrative clout in the department, which funds most US research in basic physics.

Confirmation of Moniz's appointment by the Senate may take some time, even though it is generally regarded as a formality. But the nomination has led to speculation that Martha Krebs, assistant secretary for energy research and the most senior scientific official in the department, may leave. Krebs and Moniz get along personally, says one scientist, but have clashed on technical issues.

'First trial' of anti-HIV gene therapy

[SYDNEY] Australian scientists have begun what they claim to be the world's first trial of a gene therapy for inhibiting HIV infection. The phase I trial, on male twins, is aimed initially at testing the safety of the use of ribozymes in humans. It is being run by a team from two Sydney-based companies — Gene Shears and Johnson & Johnson Research — and St Vincent's Hospital.

This is the first clinical application of technology developed from research carried out by the Commonwealth Scientific and Industrial Research Organisation — a partner in Gene Shears — and published in 1988 by researchers Wayne Gerlach, now with Johnson & Johnson, and Jim Haseloff, now at the University of Cambridge (see *Nature* 334, 585; 1988 & 342, 609; 1989).

Court awards damages over pesticide exposure

[LONDON] Campaigners preaching the potential health risks of organophosphate (OP) pesticides received a boost last week when a Hong Kong court awarded £1.9 million (US\$3.1 million) to an American musician who suffered health damage after OP pesticides were sprayed in a concert hall 10 years ago.

The action centred on diazinon, made under licence from Ciba-Geigy, one of the defendants. This chemical is suspected of links to 'Gulf War syndrome' and of causing health problems in farmers who use it to dip sheep. Elizabeth Sigmund, a long-time campaigner who is conducting research on the impact of OPs on the human nervous system in association with the Glasgow Southern General Hospital, said the decision backed claims that the extent of OP poisoning has been underestimated.

Cloning, dignity and ethical reasoning

Sir — Axel Kahn complains (*Nature* 388, 320; 1997) that his argument that “application to man of asexual reproduction and cloning represents an affront to human dignity” has been misreported by *Nature* and by John Harris’s letter. He usefully restates his argument. This remains confused, like the all-too-influential French national committee report, which he helped to draft. It is important to try to cast light on the confusion.

Kahn ascribes autonomy to “the indeterminability of the individual with respect to external human will”. He then asserts that “asexual reproduction would lead to... people whose bodily form and genetic make-up would be exactly as decided by other humans”. This is an interesting example of the genetic determinism so rightly condemned by the French national committee report. Could even so distinguished a molecular biologist as Kahn actually predict in advance the adult height and body weight of any given clone? And, even if he could, why should that affect the autonomy of the clone’s behaviour? After all, he rightly argues that the genetic make-up of a sexually reproduced person does not affect the autonomy of that person’s behaviour.

Kahn then asserts that the relationship between ‘creator’ and ‘created’ would be “entirely new”. But why, and in what respects? He further asserts that this “has obvious implications for human dignity”.

Again, why? At a Unesco meeting in May, another contributor to the French national committee report argued that cloning might lead to the creation of second-class citizens and even to a revival of slavery. The answer to that rhetoric is that humans, alas, have been well capable of creating second-class citizens and slaves without any resort to the technology of cloning.

Could we now have a philosophically serious debate about the only morally appropriate proposed use of human cloning, namely Robert Winston’s suggestion (*Br. Med. J.* 314, 913–914; 1997) that it might provide reproductive possibilities “for those men who exhibit total germ cell failure”? Winston points out that “even if straight cloning techniques were used, the mother would contribute important constituents — her mitochondrial genes, intrauterine influences, and subsequent nurture”.

Let us make two assumptions: first, that in any society where this was thought appropriate, such cloning would be tightly regulated. Second, that many years of animal trials would have established that human clones were not likely to be at undue physiological risks compared with sexually reproduced humans. At this point, many would argue that the animal experiments would not be justified. If that argument were accepted, then, on those grounds alone, it would be right to rule out human cloning.

If, however, the necessary animal trials

had been conducted, then what objections might be made to the limited application of cloning to overcome extreme cases of infertility? Kahn asserts, first, that the parental relationship would be “entirely new”. Not so. We have many models of parenting: adoption, fostering, step relationships.... All may draw on the model of a two-parent nuclear family — in itself a relatively modern phenomenon. So, with cloning, there might be new aspects to the parental relationship. To argue that it would be “entirely new” is a piece of unjustified rhetoric: studies of new reproductive techniques suggest that families have little difficulty in assimilating them into traditional patterns (see, for example, J. Edwards *et al. Technologies of Procreation: Kinship in The Age of Assisted Procreation*, Manchester Univ. Press, 1993).

The second argument is that “an individual must never be used exclusively as a means”. Kahn makes great play of the force of “exclusively”. Perhaps he could explain why the cloning of a child for a man with total germ cell failure results in that child being used “exclusively as a means”. How does that case differ from that of children resulting from medically assisted reproduction? The latter accounted in the early 1990s for almost 12 per cent of all births in France.

David Shapiro

*Nuffield Council on Bioethics,
28 Bedford Square,
London WC1B 3EG, UK*

Back to basics

Sir — Misleading references in scientific papers are common where a research group has studied the same topic for many years.

I recently gave a course on the energy metabolism of *Escherichia coli*. In a paper published in 1996, the authors say: “For β -galactosidase assay, cells were grown in a glucose-containing minimal medium (pH 7.0) unless otherwise indicated”⁴.

Reference 4 is a paper from the same research group published in 1990. In this paper the sentence reads: “For the β -galactosidase assay, cells were grown in glucose (40 mM) minimal medium (pH 7.0) (ref. 13) unless otherwise indicated”.

Reference 13, published in 1985, reads: “For galactokinase assays and mRNA isolation, cells were grown in 50-ml volumes of glucose minimal medium”³² supplemented with the indicated electron acceptor”.

Finally, the composition of the

medium is given in reference 32.

This is a random example of a very bad habit. But why don’t authors give the correct reference in every paper?

Jukka Heinonen

*Department of Biochemistry,
University of Turku,
FIN-20014 Turku, Finland
e-mail: jukahe@utu.fi*

Alphabetical orders

Sir — Recent correspondents^{1,2} have suggested that authors with initials earlier in the alphabet are more likely to be cited simply because they are more numerous, a fact borne out by examination of the distribution of surnames in the London telephone book.

However, comparing the citation rate of each letter with the number of telephone users of each letter is an inferior attempt to control for differences in the representation of initials to the

method employed in the original study.

The correlation originally identified³ was between the number of citations for each paper published by authors of a particular initial and their surname’s alphabetic position. This controls for the uneven distribution of initials among authors — if Zs are rare, they will publish fewer papers.

Perhaps the crude technique of using Londoners as a control removes the correlation because only a subset of papers are published by people with English names. Clearly there is still a case to answer, and given the growing importance of citation indices, one deserving of serious research.

Tom Tregenza

*Department of Biology,
University of Leeds,
Leeds LS2 9JT, UK
e-mail: gentbt@leeds.ac.uk*

1. Shevlin, M. & Davies, N. O. *Nature* 388, 14 (1997)

2. Jaekel, K. M. *Nature* 386, 643 (1997)

3. Tregenza, T. *Nature* 385, 480 (1997)

Turning back the harmful red tide

Harmful algal blooms are a serious and increasing problem in marine waters, yet scientists and funding agencies have been slow to investigate possible control strategies.

Donald M. Anderson

Each holiday season I await the visit of one relative with trepidation. Years ago he asked whether I had “stopped that red tide problem yet?” — a simple question from one convinced that science solves problems directed to a so-called expert on the destructive and often visible ‘blooms’ of phytoplankton that kill fish, make shellfish poisonous and cause numerous other problems in coastal waters. I explained that we did not understand the causative organisms, their ecology or oceanography well enough to propose control strategies, but that one day we would.

Although temporarily satisfied with my argument, each year thereafter my brother-in-law repeated the question, and each year my answer was the same. Whatever progress had been made, there were new questions to be addressed. Eventually, he concluded that I did not want to solve the problem, as that would end my research programme. He is wrong, of course, but the explanation is far more complex than he would think, and is in part the subject of this article.

Throughout history, blooms of microscopic algae have had a major impact on fish, birds, mammals and other organisms in the marine food web. These ‘red tides’ (now termed harmful algal blooms) take many forms and have many effects. Some toxic algae kill wild and farmed fish. Others produce potent neurotoxins that accumulate in filter-feeding shellfish and poison human consumers. Algal toxins can alter the structure and function of marine ecosystems, affecting fecundity and survival at all levels.

Even non-toxic algae can be harmful when they accumulate in sufficient numbers — sometimes millions of cells per litre — to discolour the water, shade submerged vegetation, disrupt food-web dynamics and cause oxygen depletion. At the other extreme, toxic algae can be a tiny fraction of the total phytoplankton population and still be dangerous. Diarrhetic shellfish poisoning, for example, has been reported with *Dinophysis* concentrations of a few hundred cells per litre.

The scale and timing of harmful algal blooms is highly variable. Some are localized, occurring in bays or estuaries; others are massive, covering thousands of square kilometres. Some occur at the same time and place each year; others strike in random fashion. Some last a few weeks, others years.

Harmful algal blooms are not new phenomena. Red tides are recorded in the Bible and in the fossil record. What is new is the

recent proliferation of harmful blooms¹. There is debate about the nature and causes of this expansion. Some call it a global epidemic linked to pollution and human changes to coastal ecosystems². Others argue that the expansion is in part an artefact reflecting increases in the number of scientists, advances in toxin detection, and the proliferation of aquaculture and other activities requiring product monitoring^{1,3}.

One thing is certain — there is a growing global problem at a time when human reliance on coastal zones for food, recreation and commerce is rapidly expanding. Nevertheless, there is practically no exploration of direct control of marine blooms — attempting to kill or remove the cells or reduce their toxicity. At an international conference on harmful algae held in Vigo, Spain, in June, only one contribution of more than 400 abstracts from 58 countries addressed direct control of marine blooms. Imagine the difference if the conference had been on agricultural pests or on algal blooms in fresh water, where control efforts are common.

Research efforts on mitigation strategies such as shellfish-monitoring and aquaculture site management are critically important, but they treat the symptoms without attacking the problem. Government officials and the public want to know what is being done, or what can be done, in terms of direct intervention. So far, we have little to offer other than tentative predictions of bloom reductions decades from now if nutrient loadings are reduced.

I believe that some harmful algal blooms can be controlled or managed, not 20 years from now, but in the near future, economically and without disastrous environmental consequences. This belief may brand me as a heretic among my colleagues, some of whom fear that the ocean will be further despoiled by inept human attempts to manipulate ecosystems we do not understand.

At the heart of this negativism is a conviction that mankind does not possess the skills, knowledge or right to manipulate the marine environment on any significant scale. We are, however, already doing exactly that. By polluting coastal waters, we change the abundance and relative amounts of critical plant nutrients, which in turn alters the species composition of planktonic ecosystems. Indeed, this may be why there is an increasing number of harmful algal blooms. We are harvesting fish and shellfish at an alarming rate, removing components of the food-web with little knowledge of how such enormous manipulations affect other levels.



Some red tides, such as this non-toxic bloom of *Noctiluca* off California, cover huge areas, making it difficult to foresee environmentally benign bloom-control strategies (see also <http://www.redtide.whoi.edu/hab/>).

To replace dwindling natural fishery resources, we are turning coastal waters and wetlands into marine farms at an extraordinary rate. Whether by fish or shrimp farms (which have been likened to small cities with respect to their production of organic matter as pollutants), or by shellfish or seaweed culture (which strip plankton and nutrients from the water), we are altering near-shore waters significantly. Coastal ecosystems are no longer pristine and will not revert to their ‘natural’ state without intervention.

From land to ocean

Distrust of our ability to control pests and diseases seems to be based more on pessimism than on fact. When biological control is discussed, for example, some are quick to point out failures such as the introduction of the mongoose to oceanic islands or the giant toad to Australia⁴. Obscured by these failures is a multitude of successes in terrestrial biocontrol of weeds and pests⁴. Overall, 165 insect pests and 35 weed species have been controlled. Less than 2 per cent of the introductions became pests themselves, and many of those were ‘generalist’ predators — an approach that is no longer practised.

Other examples of terrestrial management strategies include integrated pest

management, which combines biological control with chemical agents such as narrow-spectrum pesticides, and ecologically based pest management, which attempts to work with ecosystem processes in the management effort⁵. The conceptual framework for pest management on land is far advanced, and should be a valuable resource in planning the management of marine systems. Instead it is largely ignored and misunderstood.

Extrapolation from land to the ocean will admittedly be difficult, as marine and terrestrial systems differ in scale, complexity and dynamics⁶. Application of a control agent to a single crop on a parcel of land is certainly simpler than the marine equivalent, where water motions will change the distribution and abundance of target organisms and applied control agents. Control of an outbreak at one site may have no effect on blooms in later years at the same location⁶. Another difference is that the community of organisms in marine ecosystems is more diverse and complex than that in single-crop agricultural systems.

Yet another factor that has stalled progress is the tendency to generalize that all blooms are massive. One colleague argues that blooms, like tornadoes or hurricanes, can be tracked and their movements predicted, but never controlled. He may be right about the larger blooms, but many are small or localized, either permanently or during key stages of development. For example, destructive brown tides in Texas or New York are thought to begin in certain bays and then to spread to adjacent waters. A widespread coastal bloom might be localized and accessible at an earlier stage.

Another constraint is that most countries have no official policy for funding marine pest management. In the United States, the Department of Agriculture puts extensive resources into terrestrial pest management. By contrast, the equivalent agency responsible for the oceans, the National Oceanic and Atmospheric Administration, has no marine pest-control programmes. Lacking strong leadership or targeted funding programmes at the national level, scientists opt for research on fundamental ecological or oceanographic issues less likely to be rejected during peer-review.

Control options

But there are signs of change in at least some parts of the world. South Korea has established a harmful algal bloom engineering division at its National Fisheries Research and Development Institute. And Australia's Commonwealth Scientific and Industrial Research Organization has established a Centre for Research on Introduced Marine Pests, which plans a proactive approach to marine pest management consistent with the country's aggressive reliance on terres-

trial biological control. It is not clear whether this new programme will support research on control of harmful algal blooms.

During a red tide in Florida 40 years ago, copper sulphate was applied to 10,000 acres of shoreline using crop-dusting planes⁷. The treatment was initially effective, but blooms reappeared within weeks. Copper was deemed too expensive and non-specific to be used other than for short-term, small-scale bloom control. In another study, 4,700 chemicals were screened for use against Florida's red-tide alga, but none was sufficiently potent in natural sea water without also having adverse effects on other organisms.

Since then, chemical control options have received little attention, and no significant bloom-control research has been undertaken in the United States. Japan, China and Korea are exploring control strategies because they 'farm' their coastal waters heavily through aquaculture. Faced with significant economic losses from red tides, Japan initiated a broad evaluation of bloom-control strategies in the mid-1970s. Much of our knowledge of possible approaches comes from this outdated but useful series of studies⁹, which continues to this day, but at a much-reduced level of effort.

One promising strategy treats blooms with flocculents such as clay that scavenge particles, including algal cells, from sea water and carry them to bottom sediments. Removal efficiencies of 95–99 per cent have been achieved in laboratory cultures using clay, and small-scale field trials near fish farms have also been successful, though expensive⁸. New clays that are an order of magnitude more efficient in cell removal have been tested in China; this capacity can be further increased using coagulants such as polyhydroxy aluminium chloride⁹.

Applications in China have been limited to tests in shrimp aquaculture ponds but, in 1996, 60,000 tons of clay were used in Korea to control a *Cochlodinium* red tide threatening near-shore fish farms. About 100 km² were treated over several weeks, and nearly 100,000 tonnes of clay are now stockpiled in anticipation of the next bloom.

Clay is a non-toxic, naturally occurring material. Fish and bottom-dwelling organisms are unaffected by extremely high clay loadings near pottery industries¹⁰. The prospects look good, but considerable research is needed before large-scale field applications can be attempted. Obvious concerns are the fate and effects of sedimented cells and toxins on bottom-dwelling animals, and the collateral mortality of co-occurring planktonic organisms.

Biological control of harmful algal blooms also has potential. Zooplankton that graze on bloom species have been proposed as control agents⁸ but remain untested because of the impracticality of growing and maintaining predators in sufficient quantity.

Viruses are abundant in marine systems, replicate rapidly and tend to be host-specific, suggesting that a single algal species could be targeted¹¹. Parasites¹² also have potential to control algal bloom species, but specificity is largely unknown. There are numerous examples of bacterial strains¹³ exhibiting strong and specific algicidal activity, although no field applications have yet been attempted.

Prognosis for the future

I have mentioned only a few of many potential control strategies. We must cautiously explore all possible approaches, but this requires funding at the scale needed to provide data to support informed decisions and override our preconceptions. We also need to establish guidelines for acceptable marine treatments.

In one sense, the problems we face with harmful algal blooms are similar to those encountered in agriculture or medicine, fields in which control of pests and diseases is a practical reality. The marine environment is admittedly different, but our hesitancy reflects a *de facto* acceptance that the problems are insurmountable. I believe they are not, and that we can make progress if new resources are made available and if we learn from the mistakes and successes of more than 100 years of experience controlling terrestrial pests.

There is a thin line to walk here—to argue that we have been too cautious and that success is possible, without promising more than we can deliver, unrealistically raising the expectations of the public and politicians. I also worry that funds needed for ecological studies might be diverted to control. I see the risks, but I also see the chance that this article may initiate a debate that will ultimately direct scientific thought and resources towards practical solutions. My brother-in-law would no doubt approve. □

Donald M. Anderson is at the Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.

e-mail: danderson@whoi.edu

1. Anderson, D. M. in *Red Tides: Biology, Environmental Science and Toxicology* (eds Okaichi, T., Anderson, D. M. & Nemoto, T.) 11–16 (Elsevier, New York, 1989).
2. Smayda, T. J. in *Novel Phytoplankton Blooms* (eds Cosper, E. M. et al.) 213–228 (Springer, New York, 1989).
3. Hallegraeff, G. M. *Phycologia* **32**, 79–99 (1993).
4. Greathead, D. J. in *Biological Control: Benefits and Risks* (eds Hokkanen, H. M. T. & Lynch, J. M.) (Cambridge Univ. Press, 1995).
5. *Ecologically based Pest Management: New Solutions for a New Century* (National Academy, Washington, DC, 1996).
6. Lafferty, K. D. & Kuris, A. M. *Ecology* **77**, 1989–2000 (1996).
7. Rounsefell, G. A. & Evans, J. E. *US Fish Wildlife Serv. Spec. Sci. Rep.* **270** (1958).
8. Shirota, A. Int. J. *Aquacult. Fisheries Technol.* **1**, 195 (1989).
9. Yu, Z., Zou, J. Z. & Ma, X. *Oceanologia et Limnologia Sinica* **25**, 226–232 (1994).
10. Howell, B. R. & Shelton, R. G. *J. Mar. Biol. Ass. UK* **50**, 593–607 (1970).
11. Milligan, K. L. D. & Cosper, E. M. *Science* **266**, 805 (1994).
12. Taylor, F. J. R. *J. Fish. Res. Board. Can.* **25**, 2241–2245 (1968).
13. Yoshinaga, I., Kawai, T. & Ishida, Y. *Fish. Sci.* **63**, 94 (1997).

A bug with excess gastric avidity

Russell F. Doolittle

The 1,667,867-base-pair genome of the bacterium that is responsible for peptic ulcers has been completely sequenced. Among the many features revealed is machinery for existence in an acidic environment.

Yet another completely sequenced bacterial genome! The 1,667,867 base pairs of *Helicobacter pylori*, reported by Tomb *et al.*¹ on page 539 of this issue, constitute the seventh completely sequenced bacterial genome (Table 1), and the fourth by The Institute for Genome Research (TIGR). With another 30 such projects likely to be completed in various labs around the world during the next two years, some readers may be wondering when enough is enough. Is this new sequence big news? The answer has to be a resounding "yes" — *Helicobacter pylori* (a.k.a. *Campylobacter pylori*) is the organism that causes peptic ulcers and, astonishingly, it may infect as much as half of the world's population.

In 1983 an Australian physician, Robin Warren, and his colleague Barry Marshall, a biochemist, wrote separate letters to *The Lancet* under a single title². Warren described how, over the previous five years or so, he had observed spiral-shaped bacteria on the gastric epithelium of patients with peptic ulcers. Marshall then discussed the history of pyloric bacteria in mammals, and offered some thoughts on why these *Campylobacter*-like organisms had not been reported previously. He noted that they had been found in non-human mammals, and that they were even observed long ago in human cadavers, but their existence had been rationalized as a post-mortem consequence.

In 1984, the two wrote an article together in *The Lancet*³ offering the hypothesis that these bacteria were actually the cause of peptic ulcers, a very common malady attributed to stress, diet and excess gastric acidity. At about the same time, a large epidemiological study of peptic ulcers was published in the *Medical Journal of Australia*⁴. The authors ascribed the varying incidence of peptic ulcers in different Australian states to environmental factors, such as contaminated drinking water. An accompanying editorial⁵ stated that "the agent is probably not a bacterium". Warren responded immediately, stating that the agent probably was a bacterium⁶.

Interestingly, the epidemiological study was based entirely on records of state-mandated prescriptions for the antihistamine drug cimetidine (better known by its brand name Tagamet) which, according to news-

paper accounts, was the most prescribed drug in the world during the 1980s. Imagine the impact of a finding that penicillin or other antibacterial agents might be a more appropriate treatment. Henrik Ibsen would have loved it!

Bemused by the situation, I searched Medline for reports involving *H. pylori*, as well as for cimetidine and, for reasons that will become apparent, for the enzyme urease. These were plotted against five-year time periods, and the graph (Fig. 1) speaks for itself — papers dealing with *H. pylori* and the enzyme urease have greatly increased in number, whereas cimetidine citations are in precipitous decline. But why is urease suddenly so popular? The answer is that this enzyme is crucial to the survival of *H. pylori* in the very acid pH of the stomach. (Incidentally, urease is also a sentimental favourite among biochemists because, in 1926, it was the first enzyme ever to be crystallized.)

It has long been known that resident microorganisms in the stomach use the enzyme urease to convert urea to ammonia and carbon dioxide⁷. Few other organisms can survive in this acidic environment, but *H. pylori* has an electropositive internal milieu which helps it to fend off the onslaught of protons in the surrounding medium. Tomb *et al.*¹ have now shown that the proteins in *H. pylori* contain twice as many of the basic amino acids arginine and lysine as proteins in other organisms. Beyond that, the positively charged ammonium ions must contribute greatly to this effect.

Urease is the main antigenic activity associated with *H. pylori* and it is a convenient diagnostic tool⁸, which is why there has been

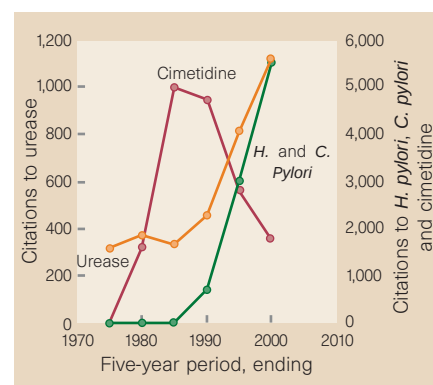


Figure 1 Citations to *Helicobacter* (or *Campylobacter*) *pylori*, cimetidine (Tagamet) and urease plotted in five-year increments. Note the abrupt changes in the period following Marshall and Warren's report³ in 1984, that *H. pylori* probably causes peptic ulcers (data for the year 2000 extrapolated from 1997).

such an enormous increase in citations involving urease. Many of the previous observations are now put into a proper context by having the entire genomic sequence. For example, all indications are that the 'pathogenicity island' is the main virulence factor. This 40-kilobase segment contains a set of accessory factors for secretion of the cytotoxins that damage the host gut⁹. The region is flanked by insertion elements and is probably the result of horizontal gene transfer from some other organism.

The peptic ulcer-urease story is, however, hardly the only reason that the sequence reported by Tomb *et al.*¹ is news. This is a genome that has something for everyone — clinicians, sociologists, epidemiologists, biochemists, ecologists, molecular biologists, immunologists and, last mentioned but hardly least interested, evolutionary biologists. Tomb *et al.* reveal the entire restriction-modification system for recognizing and destroying foreign DNA, and outline a complete scheme of metabolism on the basis of the resident genes. They have also worked out how *H. pylori* mimics blood-group antigens, as well as a likely molecular mechanism for encouraging immunogenic variation.

Of the many details reported, the generalizations about molecular evolution are

Table 1 Fully sequenced genomes

Organism	Genome size (kbp)	ORFs*	URF†	Year reported
Prokaryotes				
<i>Haemophilus influenzae</i>	1,830	1,743	732 (42%)	1995
<i>Mycoplasma genitalium</i>	580	470	187 (39%)	1995
<i>Methanococcus jannaschii</i>	1,665	1,738	1,078 (62%)	1996
<i>Mycoplasma pneumoniae</i>	816	679	344 (46%)	1996
<i>Synechocystis</i> sp.	3,570	3,168	1,766 (50%)	1996
<i>Helicobacter pylori</i>	1,668	1,590	499 (31%)	1997
Eukaryote				
<i>Saccharomyces cerevisiae</i>	12,068	5,885	2,942 (50%)	1996

* ORF (open reading frame); potential for encoding a protein.

† URF (unidentified reading frame). The numbers are approximate, many having been identified since the initial publications.

the most interesting. *H. pylori* is a Gram-negative bacterium, so most of its protein sequences would be expected to resemble other Gram-negative bacteria such as *Escherichia coli* and *Haemophilus influenzae*. Most of them do, of course, but a considerable number have sequences that are most similar to other, more distantly related, bacteria. One enzyme that is involved in chorismate biosynthesis is even reported to be most closely related to an equivalent in chloroplasts.

Although the significance of these anomalies is not clear, the possibility of rampant horizontal gene-transfer is unnerving to a community that hopes to reconstruct the history of life on the basis of amino-acid sequence comparison¹⁰. To test the point, I examined a number of bacterial urease sequences including *H. pylori* (Fig. 2). I found it unsettling not only that the generic boundaries were so ill-defined, but that the plant (jackbean) sequence was almost as similar to each of the bacterial sequences (an average of 65 per cent identity) as most of them were to one another.

There are other frustrations about the data. As in the previous studies, many of the potential gene-encoding regions have not been identified with regard to likely

function. Given that the genes for all new proteins come from other genes by duplication, in whole or in part, followed by descent with modification, I am surprised that so many open reading frames remain as unidentified reading frames. More rigorous searching will doubtless establish more relationships, as has proved to be the case in the earlier genome reports.

Nonetheless, one cannot help but be impressed by the tremendous accomplishments of Tomb *et al.*¹. Not only have they generated the data at a breathtaking pace, but the analyses are both insightful and thorough. The organizational problems associated with searching 1,600 open reading frames against existing databases,

and then making proper judgements about the findings, are awesome. Bring on the next bacterial genome sequence — don't hold that TIGR! □

Russell F. Doolittle is at the Center for Molecular Genetics, University of California at San Diego, La Jolla, California 92093-0634, USA.

1. Tomb, J.-F. *et al.* *Nature* **388**, 539–547 (1997).
2. Warren, J. R. & Marshall, B. *Lancet* **1**, 1273–1275 (1983).
3. Marshall, B. & Warren, J. R. *Lancet* **1**, 1311–1315 (1984).
4. Hugh, T. B. *et al.* *Med. J. Aust.* **141**, 81–85 (1984).
5. Piper, D. W. *Med. J. Aust.* **141**, 76 (1984).
6. Warren, J. R. *Med. J. Aust.* **141**, 477–478 (1984).
7. Kornberg, H. H. & Davies, R. E. *Physiol. Rev.* **35**, 169–177 (1955).
8. Vaira, D. *et al.* *J. Clin. Pathol.* **41**, 355–356 (1988).
9. Censini, S. *et al.* *Proc. Natl Acad. Sci. USA* **93**, 14648–14653 (1996).
10. Doolittle, R. F. *et al.* *Science* **271**, 470–476 (1996).

Asteroids

Eros's extended family

Richard P. Binzel

Lost in space and far from home. That's the fate suffered by the thousands of small asteroids wandering through the inner Solar System near the Earth, cast adrift from the main asteroid belt between Mars and Jupiter by collisions and by Jupiter's gravity. All hope of discovering their heritage would seem to be lost due to the untraceable chaotic evolution of their orbits; and although the unique spectral signatures of some near-Earth asteroids have provided links to their asteroid-belt origins^{1–3}, most near-Earth wanderers have no distinctive genealogical traits. But a new approach by Zappalà *et al.*⁴ has given new insights to the pedigree of the two largest near-Earth asteroids, 433 Eros and 1036 Ganymed (20 km and 32 km across, respectively). They find that Eros and Ganymed may be siblings, originating from a single collision that formed an entire 'family' of small asteroids, most of which still reside in the main belt.

The placement work culminating in the new paper by Zappalà and colleagues was begun seven decades ago by the Japanese astronomer K. Hirayama, who noticed that main-belt asteroids are clustered in their orbital dimensions. Hirayama called these groupings 'families' because he believed their members shared a common origin, and the collisional disruption of asteroid 'parent bodies' is now recognized as the process responsible for creating families.

One family recognized by Hirayama^{5,6} was called Maria, named after the member with the lowest catalogue number (170 Maria). Orbiting 2.55 AU from the Sun, the Maria family has about 70 identified members, and is precariously poised on the edge of one of the major gaps in the asteroid belt, a 0.05-AU-wide band of unoccupied orbits centred at 2.50 AU (1 AU is the mean Earth–Sun distance). Jupiter probably creates this

gap: a circular heliocentric orbit at 2.50 AU has a period exactly one-third of Jupiter's, and any orbit near this 1:3 resonance undergoes chaotic changes in its eccentricity⁷ (becoming more elongated). Eventually, it will have a high enough eccentricity to cross the orbits of the inner planets, so a consequence of this chaotic orbital evolution is the potential for delivering asteroids and meteorites to the inner Solar System from the main asteroid belt⁸.

A critical step in linking Eros and Ganymed to the Maria family was reconstructing the velocity field of bodies emanating from the disrupted parent body. This problem was solved using numerical techniques to determine the orbital parameters of the parent body at the time of the break-up, developed earlier by Zappalà and colleagues⁹. From this velocity field and the size distribution of the recognized family members, the authors now estimate⁴ that on the order of ten objects in the 15–30-km range should have been injected into the 1:3 resonance zone — objects about the size of Eros and Ganymed. It is not clear whether those two objects were the only ones to be successfully injected, or whether their brethren have already met their ultimate fate (collision with the Sun or a terrestrial planet, or gravitational ejection resulting from a close planetary encounter). No other asteroid families are candidates for the heritage of Eros and Ganymed. Zappalà *et al.* also compare the spectral properties of Eros and Ganymed with those of about a dozen members of the Maria family, and find agreeable matches.

Although the authors make an excellent set of plausibility arguments for the pedigree of Eros and Ganymed, a convincing case is extremely difficult. Injecting such large asteroids into a chaotic zone undoubtedly

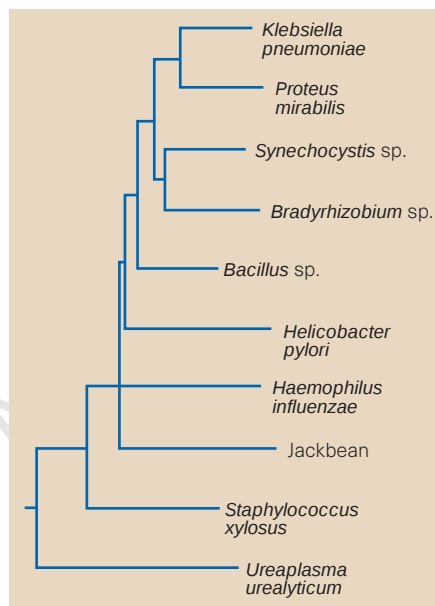


Figure 2 The genome of the Gram-negative bacterium *H. pylori*, sequenced by Tomb *et al.*¹, has thrown up a few surprises. One is that certain protein sequences do not most closely resemble those of other Gram-negative bacteria, and this may sound a warning for reconstruction of evolutionary relationships. The phylogenetic tree shown was constructed from a 480-residue, highly conserved region of the enzyme urease. On the average, the plant (jackbean) sequence is 65 per cent identical with the bacterial sequences used. The closest relationship shown is between *Klebsiella pneumoniae* and *Proteus mirabilis*, which are 77 per cent identical.



Figure 1 Two asteroids come together to make a family. This artist's impression shows colliding asteroids of different mineralogical types, which would produce meteoritic fragments of varying compositions and some breccias in which fragments of the two types are welded together into a single rock. Such a collision probably gave rise to the Maria family of asteroids, and might indirectly have sent the large near-Earth asteroids Eros and Ganymed into the inner Solar System. (Painting by William K. Hartmann.)

requires a huge collision, but that need not create a recognizable family. Furthermore, Eros and Ganymed have spectral properties that are largely similar to the majority of asteroids in the inner main belt, making their true siblings difficult to recognize.

Despite these difficulties, Zappalà *et al.* have offered an important new argument. Linking the near-Earth asteroids and their meteorite samples to their original formation locations remains one of the most important challenges for asteroid researchers today. Understanding the family history of Eros is particularly important because this large and accessible Earth-approaching asteroid has been chosen as the destination for the Near-Earth Asteroid Rendezvous (NEAR) mission, which is scheduled to arrive there in 1999. Through observations using NEAR's sophisticated sensors, Eros will become the first asteroid for which we have detailed measurements of chemical composition and elemental abundance. So if we can trace the genealogy and birthplace of Eros, these measurements will establish early Solar System conditions at a known orbital radius within the asteroid belt. □

Richard P. Binzel is in the Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

- Gaffey, M. J., Reed, K. L. & Kelley, M. S. *Icarus* **100**, 95–109 (1992).
- Binzel, R. P. & Xu, S. *Science* **260**, 186–191 (1993).
- Cruikshank, D. P., Tholen, D. J., Hartmann, W. K., Bell, J. E. & Brown, R. H. *Icarus* **89**, 1–13 (1991).
- Zappalà, V., Cellino, A., Di Martino, M., Migliorini, F. & Paolicchi, P. *Icarus* (in the press).
- Hirayama, K. *Jap. J. Astron. Geophys.* **6**, 137–162 (1928).
- Hirayama, K. *Proc. Imp. Acad. Jap.* **9**, 482–485 (1933).
- Wisdom, J. *Icarus* **56**, 51–74 (1983).
- Wisdom, J. *Nature* **315**, 731–733 (1985).
- Zappalà, V., Cellino, A., Dell'oro, A., Migliorini, F. & Paolicchi, P. *Icarus* **124**, 156–180 (1996).

Mycorrhizal fungi

The ties that bind

David Read

In most higher plants, symbiotic fungi are central to the process of nutrient capture from soil¹. Evidence from fossils of the earliest land plants², as well as molecular studies³, confirms that roots co-evolved with fungal partners to form structures known as mycorrhizas — literally, 'fungus-roots'. These are almost universally distributed through present-day terrestrial plant communities, yet most researchers (deterred, one suspects, from experimental analysis of mycorrhizal function in natural communities by the complexity of these systems) have instead used excised roots or pot-grown plants to examine the relationships between partners in the symbiosis. Unfortunately, reductionist approaches cannot answer larger questions about the effect of symbiosis on interactions between the individual plants that form natural ecosystems.

The study of Simard *et al.*⁴ (page 579 of this issue) is important in this context. Not only does it address these complex questions in a field situation but, for the first time, it shows unequivocally that considerable amounts of carbon — the energy currency of all ecosystems — can flow through the hyphae of shared fungal symbionts from tree to tree, indeed, from species to species, in a temperate forest. Because forests cover much of the land surface in the Northern Hemisphere, where they provide the main sink for atmospheric CO₂, an understanding of these aspects of their carbon economy is essential.

Hyphae are the main structural elements of mycorrhizal fungi. They either penetrate the cells of the plant root to form an 'endomycorrhiza' or, as in most of the trees in the forest studied by Simard *et al.*, they ensheath the root to produce an 'ectomycorrhiza'. In the nutrient-impovertished conditions that prevail in forests, at least 90 per cent of the 'feeding' roots of the tree are colonized by ectomycorrhizal fungi. The result is that a layer of fungal tissue, the mantle, forms an interface between the root and the soil. From this mantle, individual hyphae (~3 µm in diameter) or organized, root-like aggregates called rhizomorphs (~20 µm diameter) grow outwards to intimately explore the soil (Fig. 1). Extension of this fungal mycelium into the soil depends on a supply of photosynthetically fixed carbon from the plant¹. Conversely, essential minerals (especially nitrogen and phosphorus) captured at some distance from the root by the foraging mycelium, are transferred in the reverse direction to the root¹.

Fundamentally important for the processes investigated by Simard *et al.* is the fact that most mycorrhizal fungi are catholic

in their choice of host species⁵. As a result, the roots of trees such as the Douglas fir or birch can be colonized by many fungal species, the mycelia of which extend from tree to tree, providing linkages between them. The lack of specificity ensures that, in an undisturbed forest ecosystem, almost all of the trees — irrespective of their taxonomic affinities — are interconnected by a diverse population of mycelial systems. Groups of tree species joined together in this way have been recognized as functional guilds⁶.

Simard *et al.* provide a fascinating glimpse of one of the exchange processes facilitated by these underground connections. Using an ingenious approach in which young trees growing close to one another in the forest were simultaneously fed with either ¹⁴C- or ¹³C-labelled CO₂, they showed that net transfer of carbon occurred from birch to fir, both of which shared up to ten mutually compatible fungal symbionts. Moreover, no such transfer occurred between the species of the ectomycorrhizal guild and cedar, which was colonized by fungi of the endomycorrhizal type. Previous studies⁷ using ¹⁴CO₂ have shown that carbon transfer between plants occurs through the hyphae of compatible mycorrhizal fungi. But these studies could not confirm net flow, because transfer in the reverse



Figure 1 The roots of up to 90 per cent of the trees in a forest can be colonized by many fungal species, the mycelia of which form ectomycorrhizal networks. These networks, shown above linking two conifer species, provide channels for the transfer of nutrients. Simard *et al.*⁴ have now shown that considerable amounts of carbon can be transferred in either direction between trees through these networks.

direction was not quantified.

Perhaps of greatest ecological significance is the demonstration that when the fir was in the shade, there was a considerable increase in the amount of carbon that it received from the birch. Most forest trees spend the early part of their lives as seedlings in the gloom of the forest floor. If, as indicated by this work, the direction of flux is determined by the carbon status of the recipient, the carbon economy of this shaded understorey will be subsidized by fully illuminated overstorey plants, through pathways provided by their fungal symbionts. The extent to which this subsidy might contribute to the ability of young trees to survive in shaded environments should clearly now be examined, because persistence under, and recruitment into, the canopy determines the equilibrium of the forest community. Moreover, certain plants that live on the forest floor — notably in the family Monotropaceae — totally lack chlorophyll, so they receive all of their carbon from photosynthesizing trees through mycorrhizal connections¹. This illustrates the potential of these processes for sustaining receiver plants.

The observation that carbon is transferred between green plants will stimulate us to examine forest ecosystems from a fresh standpoint. It indicates that we should place less emphasis on competition between plants, and more on the distribution of resources within the community. If mycorrhizal colonization results in an equalization of resource availability, as suggested by this and a number of microcosm studies, it would be expected to reduce dominance of aggressive species, so promoting coexistence and greater biodiversity.

Since Clements⁸ introduced the concept of the community as a 'superorganism', ecologists have espoused the abstract idea that all components of stable ecosystems are interdependent. Simard *et al.*⁴ show the physical presence of interconnections between individuals in the forest ecosystem, and expose some of the likely consequences of their formation. The challenge now is to quantify further the contribution of these fungal linkages to the maintenance of biodiversity and stability in terrestrial ecosystems. □

David Read is in the Department of Animal and Plant Sciences, University of Sheffield, Alfred Denny Building, Sheffield S10 2TN, UK.

1. Smith, S. E. & Read, D. J. *Mycorrhizal Symbiosis* (Academic, London, 1996).
2. Remy, W., Taylor, T. N., Haas, H. & Kerp, H. *Proc. Natl Acad. Sci. USA* **91**, 11841–11843 (1994).
3. Simon, L., Bousquet, J., Levesque, R. C. & Lalonde, M. *Nature* **363**, 67–69 (1993).
4. Simard, S. W. *et al. Nature* **388**, 579–582 (1997).
5. Molina, R., Massicotte, H. & Trappe, J. in *Mycorrhizal Functioning* (ed. Allen, M. F.) 357–423 (Chapman & Hall, London, 1992).
6. Perry, D. A., Amaranthus, M. P., Borchers, J. G. & Brainerd, R. C. *BioScience* **39**, 230–237 (1989).
7. Newman, E. I. *Adv. Ecol. Res.* **18**, 243–270 (1988).
8. Clements, F. *Carnegie Inst. Wash. Publ.* **242**, 1–512 (1916).

Fluid dynamics

A fractal world of cloistered waves

Peter G. Baines

Everybody is familiar with waves on the surface of the ocean. Waves longer than about 1.7 cm are called gravity waves, because gravity (rather than surface tension) provides the principal restoring force that makes them possible. Less familiar are the similar waves that occur where there is an interface between two fluids of different densities, the lighter above the heavier. And if the density of an incompressible fluid decreases continuously upwards, waves may occur throughout the body of the fluid. Such waves are termed internal gravity waves. They permeate both the atmosphere and the ocean, and are very important for their internal dynamics on small to medium scales, where they transport momentum and cause local mixing. They may also exist in closed containers, where they converge on

'attractors' — closed curves determined by the geometry of the container — and then propagate energy along them (Fig. 1). These bizarre properties have been verified by experiments by Maas *et al.*¹, described on page 557 of this issue.

The propagation properties of internal gravity waves are much stranger than those of sound or electromagnetic waves. In particular, there is a maximum frequency that the waves may have (the buoyancy frequency); the direction in which they transmit energy (the group velocity) is perpendicular to the direction in which they propagate (the phase velocity); and in a uniformly stratified fluid, all waves of any given frequency propagate at the same angle to the horizontal, regardless of their length or structure. This last property means that if a density-strati-

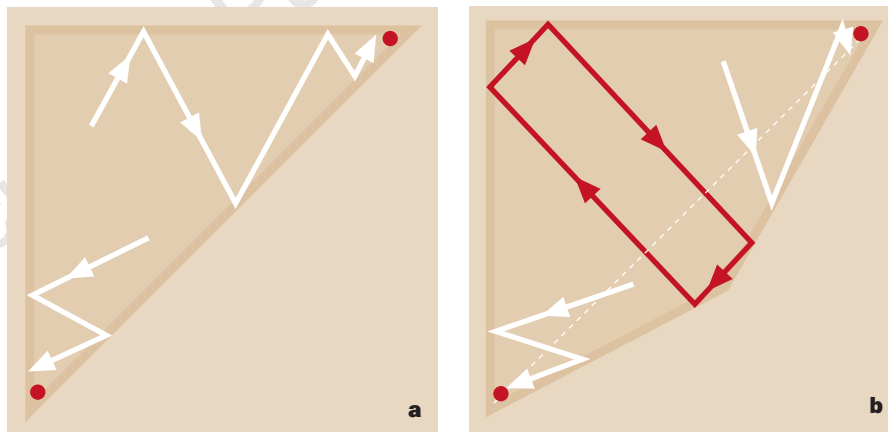


Figure 1 Models of attractor. Internal gravity waves in a fluid carry energy at an angle to the vertical that depends on their frequency. For low frequencies the rays are nearly horizontal; near the buoyancy frequency (the upper limit) the rays approach the vertical. a, In a triangular container, rays steeper than the diagonal wall go towards the right-hand vertex, and those at a shallower angle go to the bottom vertex — these two points are attractors for all internal waves. b, In a quadrilateral, rays at angles in between those of the two sloping sides converge on interior attractors. A simple example, for a particular frequency, is shown. For most of these intermediate frequencies, the attractors are much more complicated, including many circuits and having a fractal dependence on frequency.

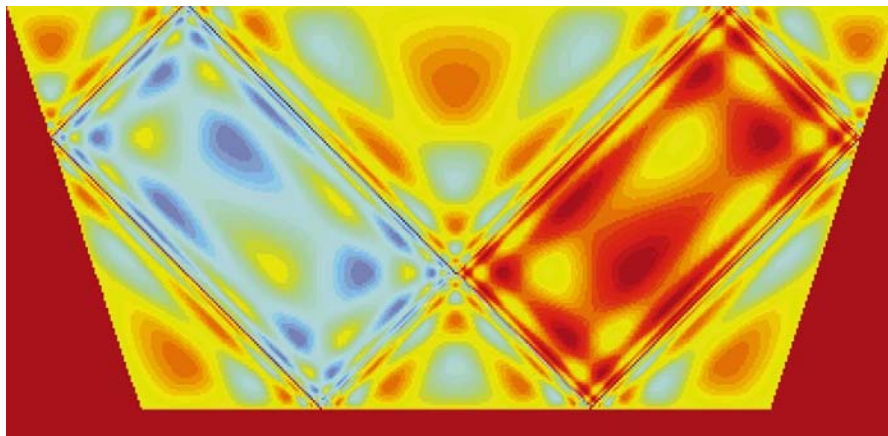


Figure 2 A wave in a bucket. For a range of wave frequencies, the fluid flow field marks out an attractor that consists of two rectangular boxes.

LEO MAAS & FRANS PETER LAM

fied fluid is forced at a frequency less than the buoyancy frequency, all the wave energy must propagate at these fixed directions. This must also apply to the waves after they reflect from solid boundaries, regardless of the orientation of the reflecting boundary. If the waves happen to be generated inside a closed container that has an irregular shape, a complex and apparently messy field of motion can result. Further, the mathematical problem of determining these properties is ill posed, and is equivalent to solving a hyperbolic partial differential equation with elliptic boundary conditions. The 'solution' is very sensitive to the geometry of the wave rays (or characteristics) inside the container. This system has an important analogue: inertial waves (due to the Coriolis force) in rotating bodies of fluid, which have similar properties.

In geophysical fluids, the principal manifestations of the unidirectional property of internal wave propagation at a particular frequency are the internal tides in the ocean. These are internal gravity waves of tidal period, generated by the oscillatory advection of the stratified ocean over topographical features on the floor of the ocean by the surface tide forced by the Sun and the Moon. Here the internal tide propagates away from the source regions into the vastness of the ocean on rays at the required vertical angle. This mathematical problem is (mostly) well posed², and these properties have been experimentally verified³ and well known for at least three decades. Field observations in the Bay of Biscay⁴ have confirmed the essential features of the wave propagation theory.

For some time, then, the basic physics of these processes in open geometries has been regarded as reasonably well known, but, in contrast, the problem of waves in irregularly shaped closed containers was thought to be too complicated and, partly in consequence, not very important or interesting. However, the growth of the theory of chaos and fractals in recent years has reawakened interest in the problem from a mathematical viewpoint. In particular, a few years ago Maas and Lam⁵ showed theoretically that, in simple containers of trapezium shape, wave energy inside the container at a single frequency may converge to an 'attractor', which consists of a closed line figure that may be as simple as a single parallelogram. The attractor is purely determined by the geometry of the container, and the wave energy converges on it and then propagates around it, as can clearly be seen in the new results¹. The shape of this attractor changes continuously with the frequency of the waves over some ranges of values, and varies fractally over others (this has similarities to the scarred wavefunctions of quantum chaos⁶).

Where such attractors occur, the motion cannot be represented as a sum (or Fourier series) of wave modes to simplify the

description of the motion, as such modes do not exist. This comes as quite a shock to most physicists, who are accustomed to representing wave motions in closed containers of all types in terms of a sum of stationary-wave modes. It is not safe to assume, for example, that internal wave motion in a given density-stratified lake can be represented in terms of a sum of modes.

The experimental confirmation now presented by Maas *et al.* shows that the simplifying concept of the attractor is relevant and useful, and it should provide a stimulus for further theoretical developments and a search for more practical applications. Possibilities for the latter include transverse wave motions in fjords, and internal tides in topographically complex regions of the ocean. There has been a resurgence of interest in internal tides, as they have recently become

detectable by satellite altimetry that measures changes in height of the sea surface, and are more prominent and ubiquitous than thought previously. But the most significant applications may lie in the rotating analogue that I mentioned earlier; here there are many possibilities, perhaps the most intriguing being the likely presence of these effects in the liquid outer core of the slightly non-spherical Earth. □

Peter G. Baines is at the CSIRO Division of Atmospheric Research, Aspendale 3195, Australia.

1. Maas, L. R. M., Benielli, D., Sommeria, J. & Lam, F.-P. A. *Nature* **388**, 557–561 (1997).
2. Baines, P. G. *Deep-Sea Res.* **29**, 307–338 (1982).
3. Mowbray, D. E. & Rarity, B. S. H. *J. Fluid Mech.* **28**, 1–16 (1967).
4. Pingree, R. D. & New, A. L. *Deep-Sea Res.* **36**, 735–758 (1989); *J. Phys. Oceanogr.* **21**, 28–39 (1991).
5. Maas, L. R. M. & Lam, F.-P. *J. Fluid Mech.* **300**, 1–41 (1995).
6. Heller, E. J. *Nature* **380**, 583–584 (1996).

Signal transduction

I κ B kinase all zipped up

Alain Israël

NF- κ B is a transcription factor that plays a central role in the immune and inflammatory response, and is also involved in controlling the expression of human immunodeficiency virus. A few years after this factor was identified¹, it became clear that its activity is controlled by a previously unknown mechanism. In most cells, NF- κ B is retained in the cytoplasm in an inactive form by the inhibitory protein I κ B. Following certain extracellular stimuli (the most frequently studied being tumour-necrosis factor (TNF), interleukin-1 (IL-1), phorbol myristate acetate (PMA) or lipopolysaccharide), NF- κ B dissociates from I κ B, translocates to the nucleus and activates its target genes (see Fig. 1, and ref. 2 for review).

Even before the first member of the NF- κ B family was cloned, it was recognized that the association of I κ B with NF- κ B is controlled by phosphorylation, and thus the search for the protein kinase responsible began. On page 548 of this issue³, DiDonato *et al.* report the purification and cloning of a cytokine-responsive I κ B kinase which might put an end to this long and frustrating search.

In 1990, Ghosh and Baltimore⁴ demonstrated that treatment of a partially purified NF- κ B/I κ B complex by certain kinases could induce dissociation of the complex and activation of NF- κ B DNA-binding activity. Cloning⁵ of the complementary DNA encoding I κ B α , one of the three forms

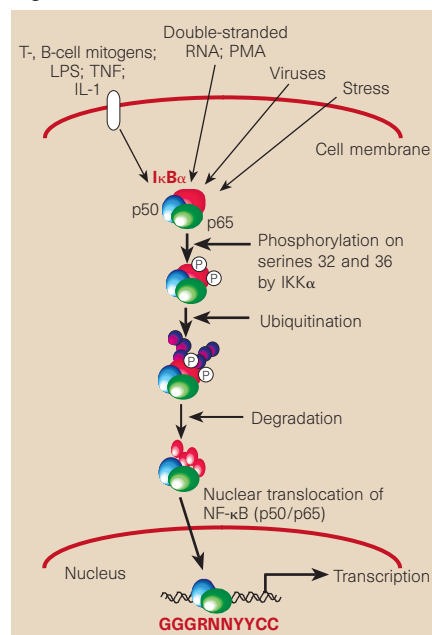


Figure 1 The NF- κ B signalling pathway (after a graphic by J. C. Epinat). The pathway is activated by various stimuli, some of which may also use alternative pathways. The stimuli include T- and B-cell mitogens, lipopolysaccharide (LPS), tumour-necrosis factor (TNF) and interleukin-1 (IL-1), double-stranded RNA or phorbol myristate acetate (PMA), infection by certain viruses, or stress (such as irradiation with ultraviolet light). The ensuing signalling cascade results in phosphorylation of the inhibitory protein, I κ B α , by IKK α , the protein kinase which DiDonato *et al.*³ have now isolated. I κ B α is then ubiquitinated and degraded by proteasomes, and NF- κ B translocates into the nucleus to modulate gene transcription. The consensus binding site for NF- κ B is shown (R, purines; Y, pyrimidines; N, any nucleotide). The five members of the NF- κ B family can form homo- or heterodimers (p50 and p65 are the constituents of NF- κ B activity). IKK α activation by stimuli other than TNF, IL-1 and PMA has not yet been investigated.

NATURE

100 YEARS AGO

It seems but the other day I saw London in a blaze of illumination to celebrate Her Majesty's happy accession to the throne. As in a few days the whole empire will be celebrating the Diamond Jubilee of our Queen, who will then have reigned over her multitudinous subjects for sixty years, what more suitable topic can I bring before you than that of diamonds! One often hears the question asked: "Why Diamond Jubilee?" I suppose it is a symbol intended to give a faint notion of the pure brilliancy and durability of the Queen's reign; and in thus associating Her Majesty with the precious diamond, to convey an idea of those noble qualities, public and private, which have earned for her the love, fealty, and reverence of her subjects. From the earliest times the diamond has occupied men's minds. ... The philosopher Steffans is accredited with the dictum that "Diamond is quartz which has arrived at self-consciousness!" and an eminent geologist has parodied this metaphysical definition, saying, "Quartz is diamond which has become insane!"

'Diamonds' – William Crookes, F.R.S.
From *Nature* 5 August 1897.

50 YEARS AGO

The "Oxford Dictionary" defines 'dialysate' as that portion of a mixture that remains after dialysis, and quotes Atfield's "Text-book of Chemistry" (1885). The quotation shows that the part that fails to pass through the membrane is referred to. From the point of view of the lexicographer, this is perfectly satisfactory; there seems to have been no earlier use of the word in print. ... Present-day usage does not agree with the "Oxford Dictionary". Nearly all the biochemists that I have consulted have understood by dialysate the more diffusible part of the system. At present, therefore, the word is ambiguous unless the context makes the meaning clear. Many authors and editors have not read Atfield, and some of those who have may question his authority in the matter. The popular meaning has probably been adopted because of the analogy with filtrate, distillate, sublimate, etc.; in each case the mobile part of the system is referred to. The continuing existence of this analogy will make it difficult to get universal acceptance of the dictionary meaning of dialysate.

From *Nature* 9 August 1947.

of IκB now known, then stimulated an intense burst of research which to some extent clarified the signalling events. In response to extracellular stimuli, the IκBα molecule becomes phosphorylated by a serine-specific kinase at positions 32 and 36. This phosphorylation in turn induces polyubiquitination of the molecule, which targets it for rapid degradation (and not for dissociation from NF-κB, as originally postulated)².

During the same period, a search for the kinase or kinases responsible for the inducible phosphorylation of IκBα led to a flurry of papers describing the direct or indirect involvement of a number of known kinases in this phosphorylation event. A bona fide IκBα kinase would be expected to specifically phosphorylate serines 32 and 36 of IκBα and to be inducible by the stimuli known to activate NF-κB; unfortunately, none of the postulated kinases fulfilled all these requirements.

This is not the case for the kinase described by DiDonato and colleagues³; they have fractionated extracts from HeLa cells stimulated with TNF, and have examined the various fractions for their ability to specifically phosphorylate a fragment of the IκBα molecule containing serines 32 and 36. The authors observed that TNF stimulation of the cells resulted in the rapid activation of a kinase activity with the expected specificity, which eluted from a gel filtration column as a complex of relative molecular mass 900,000 (900K).

Last year, T. Maniatis's group⁶ described the partial purification of a 700K complex from non-stimulated HeLa cells which contained a kinase activity able to specifically phosphorylate IκBα on the two critical serines. This kinase was unique in requiring ubiquitination in order to be activated. The same group have since demonstrated⁷ that mitogen-activated protein kinase/ERK kinase 1 (MEKK-1), a kinase involved in several signalling pathways and known to be activated by TNF, could directly activate the IκBα kinase activity of the 700K complex *in vitro*, and induce the site-specific phosphorylation of IκBα *in vivo*.

DiDonato *et al.* have purified and cloned the IκBα-specific kinase of their 900K complex, which turned out to be almost identical to a putative serine/threonine kinase of unknown function named CHUK⁸. The protein is composed of 744 amino acids, with a kinase domain in its amino-terminal region and several protein-interaction motifs, including a leucine zipper and a helix–turn–helix motif in its carboxy-terminal part. This kinase can specifically phosphorylate the two serines of IκBα and has been renamed IKKα.

IKKα can also phosphorylate the two critical serine residues of a second IκB

species, IκBβ (ref. 9). Its activity is induced by TNF, IL-1 and to a lesser extent by PMA. Transfection experiments confirmed that IKKα overexpression can activate an NF-κB-dependent reporter gene, and that an antisense RNA can inhibit NF-κB activation by TNF, IL-1 or PMA. More interestingly, DiDonato *et al.* found that IKKα activity is sensitive to dephosphorylation by the protein phosphatase PP2A, consistent with the previous observation that NF-κB activity could be induced by okadaic acid, an inhibitor of PP2A, and suggesting that, *in vivo*, IKKα is negatively regulated by PP2A or a similar phosphatase. This in turn leads to the question of the identity of the kinase or kinases responsible for the phosphorylation and activation of IKKα.

So is that it? One obvious question concerns the identity of the other proteins in the 900K complex: although NF-κB/IκB is not present, this complex might contain some of the proteins involved in ubiquitination or in the process of degradation itself (or both). Alternatively, as has been seen in yeast¹⁰, the entire kinase cascade leading to IκBα phosphorylation (or part of it) could be packed in a large complex including adaptor or scaffolding proteins, therefore ensuring speed and accuracy in the transmission of the signal (but maybe leaving less possibility for branching points).

Another question is whether IKKα is the unique integrator of the NF-κB response: in other words, do all the seemingly disparate stimuli known to activate NF-κB lead to the activation of IKKα, which in turn would phosphorylate the three known IκB species, IκBα, β and ε? If they do, then this kinase would be the perfect target for drugs aimed at controlling the inflammatory response. The report¹¹ that the mitogen-activated ribosomal S6 protein kinase, p90^{rsk}, can interact with IκBα and phosphorylate the protein upon serine 32 (but not serine 36), both *in vitro* and *in vivo* (p90^{rsk} is induced by PMA, but not by TNF), indicates that the situation might be more complicated and that different stimuli might involve different kinases. Besides, the kinases in a certain number of signalling pathways known to date seem to exist as families, and it is quite possible that IKKα also belongs to such a group of related kinases, with possibly different specificities.

Finally, some discrepancies between the new work and that of the Maniatis group need to be resolved. The 700K complex which Maniatis and colleagues isolated from non-stimulated cells can be activated either by treatment with MEKK-1 or by ubiquitination; DiDonato *et al.* did not observe a requirement for ubiquitination and were unable to activate IKK *in vitro* with MEKK-1. Does this imply that the complexes purified by the two groups are different or differently regulated, while both are relevant to IκBα

phosphorylation? The cloning of the I κ B α kinase present in the 700K complex should resolve this issue. □

Alain Israël is in the Unité de Biologie Moléculaire de l'Expression Génique, UMR 321 CNRS, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris Cedex 15, France.

e-mail: aisrael@pasteur.fr

1. Sen, R. & Baltimore, D. *Cell* **47**, 921–928 (1986).
2. Baldwin, A. S. *Annu. Rev. Immunol.* **14**, 649–683 (1996).
3. DiDonato, J. A., Hayakawa, M., Rothwarf, D. M., Zandi, E. & Karin, M. *Nature* **388**, 548–554 (1997).

4. Ghosh, S. & Baltimore, D. *Nature* **344**, 678–682 (1990).
5. Haskill, S. *et al. Cell* **65**, 1281–1289 (1991).
6. Chen, Z. J., Parent, L. & Maniatis, T. *Cell* **84**, 853–862 (1996).
7. Lee, F. S., Hagler, J., Chen, Z. J. & Maniatis, T. *Cell* **88**, 213–222 (1997).
8. Connelly, M. A. & Marcu, K. B. *Cell. Mol. Biol. Res.* **41**, 537–549 (1995).
9. Thompson, J. E., Phillips, R. J., Erdjument-Bromage, H., Tempst, P. & Ghosh, S. *Cell* **80**, 573–582 (1995).
10. Choi, K. Y., Satterberg, B., Lyons, D. M. & Elion, E. A. *Cell* **78**, 499–512 (1994).
11. Schouten, G. J. *et al. EMBO J.* **16**, 3133–3144 (1997).

Climate change

Is the ocean at the helm?

Mike McCartney

The global climate heads off, unpredictably, into different regimes of behaviour for periods of years and decades. At middle latitudes, the central question is whether the ocean is a passive participant in climate change, merely responding to the atmosphere conditions above its surface, or a far more active component of the climate. On page 563 of this issue¹, Sutton and Allen present compelling evidence that the North Atlantic is an active participant in the decadal evolving regional climate, with vivid images of propagating and oscillating sea surface temperature (SST) distributions.

Sea surface temperature distributions are a critical link between the circulations of the atmosphere and the ocean. But the physics of this link is difficult. The question is whether the atmosphere and ocean circulations are truly coupled (rather than the atmosphere directing all the action), and if so, how. Sutton and Allen have analysed patterns of wintertime SST measured in the North Atlantic over several decades, together with records of the overlying atmospheric circulation, mapped through sea-level pressure (SLP) distributions. They find that the patterns evolve in a remarkably organized way over decade timescales, propagating with slow but steady persistence in the direction of the main ocean currents. Ocean circulation is ponderous compared with that of the atmosphere, but the huge heat capacity of water gains the ocean equal partnership in the movement of heat in the climate system.

The longest continuously measured property of the oceans is surface temperature. SST was once measured by ships' crews using insulated sampling buckets, more recently by thermometers fixed in intake pipes for engine cooling water. Winter SST anomalies (departures from monthly mean values) are the surface expression of temperature anomalies in the underlying bulk of the ocean's mixed layer: the cold and windy mid-to high-latitude winter atmosphere causes convective overturning of the water column down to at least a few hundred metres, and

sometimes to more than 1,000 m. So an SST anomaly of just 1 °C represents a large amount of heat.

These heat anomalies are sequestered beneath a seasonally heated layer in summer, and re-exposed to the atmosphere in subsequent winters. This gives the ocean a 'memory' of atmospheric conditions. The repeated winter re-exposure of such accumulated thermal anomalies means that the atmosphere sees the anomaly for many winters — and there's the crux of the coupling question. If there is a feedback from anomalies in oceanic SST and heat content to the atmosphere, then recurring anomalies in atmospheric circulation should result. Is this borne out by the evidence in the North Atlantic?

One of the most robust modes of atmospheric variability is the North Atlantic Oscillation (NAO), a see-saw in the amplitudes of two centres of extreme pressure, the Icelandic low and Azorian high², which deter-

mines the strength of the westerly winds. The oscillation varies on timescales ranging from seasons to decades, and there are associated shifts in North American, European and northern Asian patterns of rainfall³ and storm tracks^{2,4}, and also temperature⁵ — which has a considerable effect on the measured pattern of 'global warming'⁶.

Figure 1 shows the winter North Atlantic Oscillation³ over 1864–1997. Are the longer timescales that predominate in this record a signature of ocean feedback to the atmosphere? Some numerical simulations have indicated that the atmosphere can oscillate decadal despite fixed SST conditions^{7,8}, so perhaps ocean memory need not be invoked to explain such low-frequency climate oscillations. But if the cause is an internal atmospheric oscillation, it must be a rather specialized one: J. W. Hurrell and I (in work in preparation) have found that, to accord with the observed changes in the pattern of SLP, the atmospheric oscillation would need to have a strong recurrent winter signal that fades in spring and is completely forgotten in the subsequent summer and autumn, and, moreover, has no precursor in the preceding autumn. It seems more likely that long-lasting SST anomalies do indeed feed back to force such atmospheric circulation anomalies on decade timescales.

The SST anomalies in question⁹, now analysed by Sutton and Allen¹, propagate along with major ocean currents and circulation pathways, so that the sequestered heat anomalies are displaced downstream each time they are re-exposed to the atmosphere. This propagation is slow, but remarkably persistent. Sutton and Allen document a nine-year majestic march of SST anomalies, starting off the Carolinas of North America

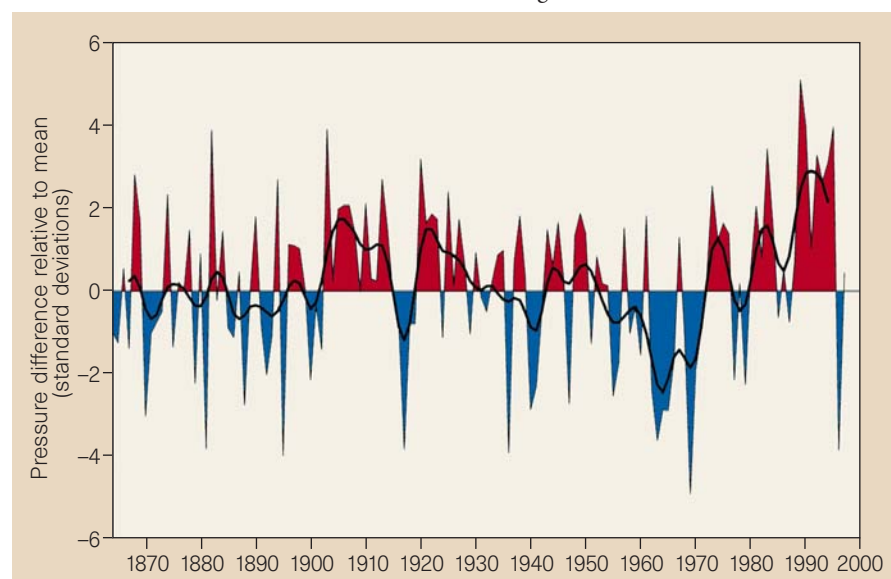


Figure 1 Winter North Atlantic Oscillation index, based on the difference of normalized sea-level pressures (SLP) between Lisbon and Stykkisholmur (in Iceland) (ref. 3 and J. W. Hurrell, personal communication). The heavy curve is the same record smoothed with a low-pass filter to remove fluctuations with periods of less than four years. This long-term variability in the westerly winds may be driven by temperature variations in the ocean.

and moving to southeast of Iceland. It would take a very contrived and special sequence of atmospheric conditions to produce air–sea heat-exchange anomalies that would yield this pattern. Comparing the NAO index of westerly wind strength (Fig. 1) with Sutton and Allen's time series of SST anomalies (Fig. 2a on page 565), I am struck by the remarkable correspondence of the period of declining NAO index from the 1940s until 1970 with the strongest propagating warm anomaly; and of the period of rising NAO index from 1970 to 1995 with the strongest propagating cold anomaly.

The SST evidence does have a somewhat curious aspect: the propagation speed of the SST anomalies, a few centimetres per second, is much slower than the principal currents involved, the Gulf Stream and North Atlantic Current. Three things contribute to that sluggishness. First, the SST anomalies are considerably wider than these currents; they span areas south of the currents where parts of the subtropical gyre are much slower and sometimes even flow westward rather than eastward, in compact recirculation cells adjacent to the main flow¹⁰. Second, because the strong currents meander over weeks to months, water parcels travelling in the current do not necessarily travel far eastward before being exchanged for other parcels to the south of the currents¹¹, and it is along such trajectories that parcels exchange heat with the air. Finally, water parcel speed and temperature propagation speed are not the same thing. Downstream cooling of the flow causes isotherms to be 'left behind'. Upstream or downstream isotherm displacements, and thus SST anomalies, could result from a fixed flow of water encountering increased or diminished cooling by the atmosphere, or from diminished or increased rate of flow of water with fixed atmospheric cooling, or a combination of both.

Climatologists have tended to think of the North Atlantic Oscillation as a mid-latitude-only phenomenon, disconnected from tropical Atlantic SST and climate change. Sutton and Allen's analysis shows wide-ranging SST anomalies that seem organized from Equator to subpolar sea. SST anomalies off the Carolinas of North America in the western subtropical gyre—the mid-latitude 'storm-formation region', as Sutton and Allen call it—appear to give rise to the slow-moving SST anomalies arriving in the eastern subpolar gyre about a decade later, and are correlated with the appearance of the same sign of anomalies in the tropical Atlantic after about six years. This has two important ramifications.

First, it shows that the SST anomalies and linked SLP anomalies are involved in a whole-ocean oscillation, a result anticipated in earlier studies^{12,13}. Second, the linking of subtropical and tropical North Atlantic SSTs

may be of particular relevance to climate forecasting. Warm conditions in the tropical North Atlantic since 1995 have been linked to the production of more tropical storms and hurricanes¹⁴, and earlier work ties the same area to decadal changes in rainfall in north Africa and northeastern Brazil¹⁵.

The demonstration of covarying SST and SLP fluctuations is evocative of a whole-ocean phenomenon, a North Atlantic atmosphere–ocean oscillation, against which attempts to simulate coupled atmosphere–ocean oscillations will be compared. Can a signal be extracted from observations that will finally confirm or deny the existence of ocean-to-atmosphere coupling? More tantalizingly, will the orderly behaviour of SST fluctuations help us to predict the aspects of climate over land that we really

care about?

Mike McCartney is in the Department of Physical Oceanography, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.

e-mail: mmccartney@whoi.edu

1. Sutton, R. T. & Allen, M. R. *Nature* **388**, 563–567 (1997).
2. Rogers, J. C. *J. Clim.* **3**, 1364–1379 (1990).
3. Hurrell, J. W. *Science* **269**, 676–679 (1995).
4. Carleton, A. M. *J. Clim.* **1**, 212–223 (1988).
5. van Loon, H. & Rogers, J. C. *Mon. Weath. Rev.* **106**, 296–310 (1978).
6. Hurrell, J. W. *Geophys. Res. Lett.* **23**, 665–668 (1996).
7. James, I. N. & James, P. M. *Nature* **342**, 53–55 (1989).
8. Marshall, J. & Moltani, E. *J. Atmos. Sci.* **50**, 1792–1818 (1993).
9. Hansen, D. V. & Bezdek, H. F. *J. Geophys. Res.* **101**, 8749–8758 (1996).
10. Schmitz, W. J. & McCartney, M. S. *Rev. Geophys.* **31**, 29–49 (1993).
11. Lozier, M. S. *et al.* *J. Phys. Oceanogr.* (in the press).
12. Barnett, T. P. *J. Atmos. Sci.* **42**, 478–501 (1985).
13. Kushnir, Y. *J. Clim.* **7**, 141–157 (1994).
14. Saunders, M. A. & Harris, A. R. *Geophys. Res. Lett.* **24**,

Carbon cycle

Inside the black box

Richard Norby

In an atmosphere in which levels of CO₂ are much higher than today, will terrestrial ecosystems accumulate more carbon or will they just process it faster? The answer to this apparently simple question has profound implications for the issue of how far the terrestrial biosphere can mop up carbon emitted to the atmosphere as CO₂ from fossil-fuel combustion. On page 576 of this issue¹, Hungate *et al.* conclude that, in two grasslands exposed to CO₂ at twice the ambient concentration, carbon cycles through the systems faster, but accumulation increases little. Consistent with model projections², the experimental results indicate that the net increase in carbon uptake observed in the short term probably does not reflect the potential for long-term carbon storage in grasslands.

The experiment was conducted on two different soil types at the Jasper Ridge

Biological Preserve in northern California. After three years' exposure to ambient or twice-ambient CO₂, the CO₂-enriched systems contained significantly more carbon, primarily in roots, detritus (dead plant material) and soil microorganisms. But total soil carbon did not increase.

Could the transfer of carbon below ground eventually lead to a long-term increase in soil organic matter, or was the carbon mostly transferred to labile, short-lived pools? Using a mass-balance analysis of the annual carbon fluxes, Hungate *et al.* found that higher levels of CO₂ primarily stimulated the annual losses of carbon through root respiration, turnover and exudation of organic carbon. Direct observations of root turnover using miniature cameras in tubes inserted into the soil, and an independent assessment of respiration in a ¹³C-labelled microcosm, suggested that the



Figure 1 The Jasper Ridge CO₂ experiment, carried out by Hungate *et al.*¹, showing some of the field chambers on serpentine soil. (Photo by J. Canadell.)

most responsive fluxes were root respiration and exudation, which have little capacity for increasing carbon accumulation in the system.

This study, and similar attempts to understand the carbon balance of different ecosystems under current and future atmospheric conditions, are a key part of the endeavour to understand global change. A central, defining objective has been to balance the global carbon cycle or, as it is often stated, to "find the missing carbon". Only about 45 per cent of the carbon released through human activities (fossil-fuel combustion, cement manufacturing and deforestation) remains in the atmosphere³. Some of the remaining fraction is absorbed by the oceans; the rest is often assigned to the terrestrial biosphere, which is treated as a 'black box'.

The challenge to terrestrial ecologists has been to find that missing carbon, or at least to better describe the fluxes into and out of the biospheric black box. Most plant systems have increased rates of photosynthesis and productivity when grown under high-CO₂ conditions. This will affect the global carbon budget only if the carbon taken up in photosynthesis is sequestered; that is, if the increased influx is not matched by a similar increase in the efflux of carbon from the system through plant or soil respiration. Accounting for all the fluxes and documenting carbon sequestration, however, has proved to be very difficult⁴. Typically, any fraction of the net influx of carbon to a system that cannot be accounted for in plant biomass is simply assumed to have been allocated below ground.

But creating this new black box is very unsatisfactory. Carbon allocated below ground could be respired by roots and immediately lost from the system. Or it could contribute to increased root growth, either of short-lived roots that quickly decompose or, in woody perennial systems, of longer-lived roots that are part of the system's carbon storage. Some carbon may exude from or be excreted by roots as low-molecular-weight organic compounds, which are subsequently metabolized by soil microorganisms. Particulate soil organic matter can accumulate if it is sheltered from further decomposition by being physically protected in aggregates⁵. Some of the carbon entering the soil system may eventually end up as recalcitrant soil organic matter and contribute to a long-lived carbon storage pool. This flux is expected to be small and is very difficult to measure, but it may increase as the overall cycling rate increases.

We cannot judge whether an ecosystem is sequestering or cycling the additional carbon without knowing which of these possible pathways predominates. Hungate *et al.* were able to peer inside this below-ground black box by taking advantage of the

annual turnover in their system and by incorporating an impressive array of complementary observations into their analysis.

Although these new results cast doubt on the capacity of the California grasslands to sequester much extra carbon, they cannot necessarily be broadly applied to other ecosystems. In particular, forests can store more carbon than grasslands. They sequester a substantial amount of carbon in woody stems and roots, a carbon pool completely absent from grasslands. Moreover, it appears that root exudation is quantitatively much less important in forests than in grasslands, and may not be a significant route for additional carbon loss in conditions of higher CO₂ (ref. 6). Leaf litter from trees may be slower to decompose and contain a larger recalcitrant fraction than grassland detritus. Attempts to construct carbon budgets for CO₂-enriched woody systems have been generally unsuccessful⁴ — in future work to that end, it is especially important that the integrated, multifaceted approach of Hungate *et al.* should be employed.

The research emphasis on carbon pools and fluxes, and on the potential for ecosystems to sequester more carbon, stems from policy-driven questions and easily defined challenges such as "finding the missing carbon". But even as we follow the reductionist approach to measuring a gradually shrinking black box, a broader, system-level view is also required. The soil system is incredibly complex, with uncounted bacterial, fungal and microfaunal species living and interacting amidst a matrix of plant roots and organic and inorganic particles, and awash in a nutrient and organic bath. Even if increased CO₂ does not lead directly to carbon accumulation, a faster cycling rate could induce myriad changes in species diversity and functions. These fundamental shifts in ecosystem physiology could in the long run be the most important controllers of carbon pools. □

Richard Norby is at the Environmental Sciences Division of Oak Ridge National Laboratory, Building 1059, PO Box 2008, Oak Ridge, Tennessee 37831-6422, USA.

1. Hungate, B. A. *et al.* *Nature* **388**, 576–579 (1997).
2. Parton, W. J. *et al.* *Glob. Change Biol.* **1**, 13–22 (1995).
3. Schimel, D. S. *Glob. Change Biol.* **1**, 77–91 (1995).
4. Canadell, J. G., Pitelka, L. F. & Ingram, J. S. I. *Plant Soil* **187**, 391–400 (1996).
5. Jastrow, J. D. *Soil Biol. Biochem.* **28**, 665–676 (1996).
6. Norby, R. J., O'Neill, E. G., Hood, W. G. & Luxmoore, R. J. *Tree Physiol.* **3**, 203–210 (1987).

Erratum

In the 10 July issue, an incorrect caption was given to the News and Views article 'How one Galaxy can be a cluster' by Richard Mushotzky (Vol. 388, pages 126–127). The caption referred to the 'dark' X-ray cluster MG2016, whereas the image shown was of a nearby, normal X-ray cluster, intended by the author to contrast with MG2016.

Daedalus

Counting on the truth

If the first major biological invention was communication, the very next was the lie. Lies are everywhere; from the harmless flies who imitate stinging wasps, to the forgers and confidence tricksters of human crime. One defence is to make the truth too costly or difficult to imitate. A big male toad can exploit his large resonant volume to make a deep, seductive croak, which a smaller, less desirable male cannot imitate. Similarly, a dominant communicator such as *Nature* can command belief by the costly presses and editorial staff needed for a printed journal. A lone scientific crank cannot afford such conspicuous authority.

But cheap electronic communication is changing all this. Any nerd with a modem and computer can have his own Web site, as seemingly authoritative as *Nature's* own. Furthermore, computers simply invite mischief-makers, as proved by the hacking and computer-virus industries. Already the Internet is dense with lies and nonsense. On tomorrow's information superhighway, says Daedalus, all the frauds and deceptions which can be spread, will be spread. Tricksters will infiltrate and subvert all communications. They may falsify the *Nature* Web site, or copy it, insert some mad paper of their own, and present the result as a *Nature* mirror site. With old-fashioned print a distant memory, how will *Nature*, or anyone else, retain their integrity? If they cannot, nobody in the glorious electronic future will have any reason to believe anything they read.

In this connection, Daedalus recalls the self-referential problem of writing a true sentence of the form, 'This sentence contains eighteen 'a's, five 'b's ...' and on down the alphabet. It is almost impossible, because the letters in the names of the numbers are also part of the task. So, says Daedalus, let every electronic issue of *Nature* contain a sentence giving its total alphabetic count, including the letters in the check sentence itself. *Nature* will need a vastly expensive supercomputer to work it out in the one week available for each issue. Any reader will be able to validate the count. But the smallest change will wreck it, and no forger could afford the computer time to derive a new matching correct sentence.

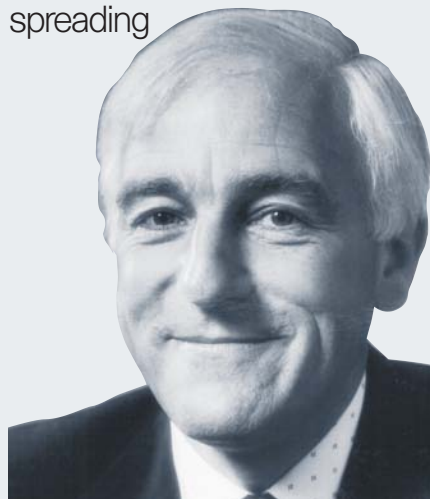
This strategy will restore credibility to dominant, respectable publishers with reputations to defend, even in the coming electronic age. The mass of little forgers and liars waiting to infest the information superhighway will be frustrated. Only big, rich liars will survive; and we're used to them already.

David Jones

Obituary

Drummond Hoyle Matthews (1931–97)

Marine geologist who discovered seafloor spreading



The plate-tectonics revolution of the late 1960s was one of the most fundamental changes in the way that we view the geology of the Earth: at a stroke it turned our thoughts from local up-and-down motions of mountains and geosynclines (sedimentary basins), to the astonishingly rapid (on geological and even human timescales) horizontal motions of the lithospheric plates that continually create and destroy vast ocean basins, volcanic chains and mountain belts. Drummond Matthews, who died of a heart attack on 20 July, lived through — and was himself one of the chief architects of — this quantum leap in our understanding of the world.

After obligatory national service (in his case, in the Royal Navy), he took a degree in geology and petrology at Cambridge in 1955. He then spent a couple of years as a geologist in the Antarctic, with the fore-runner of the British Antarctic Survey, before returning to Cambridge to complete a Ph.D. on rocks dredged from the North Atlantic.

The work for which Matthews is perhaps best known — the identification of seafloor magnetic stripes as being due to oceanic spreading — was done with his first research student, Fred Vine, and published in 1963. He discovered the seafloor spreading stripes in the Gulf of Aden during cruises on HMS *Owen*, as part of the international Indian Ocean expedition of 1961–63. On board ship, he worked without technical support, pulling in by hand the 500-metre towed magnetometer cable, as well as running the instrumentation.

On the untimely death of his mentor Maurice Hill, Matthews took over the running of the Cambridge marine geophysics group. Matthews once called this time “the heroic period of ocean exploration” — and heroic it certainly was. In the five years from 1967 to 1972, the members of the 15-strong marine group took part in 33 sea-going expeditions. Almost every cruise was to virgin territory, and they discovered many new features of the ocean floor and its plate boundaries.

The pace of exploration and innovation scarcely slackened during the next decade: during 1972–82, the Cambridge marine group ran another 39 expeditions at sea. Taking advantage (in the days before the delimitation of vast offshore areas as ‘exclusive economic zones’) of being able to sail a research ship almost anywhere, Matthews addressed the most tractable problems in the most favourable locations. His laboratory was the two-thirds of the world that is covered by water.

Under Matthews’ leadership, the marine group continued at the cutting edge of international science. His breadth of vision is illustrated by one study that he led, of why sedimentary basins (such as the North Sea) subside and accumulate thick sediment sequences. In 1978, Dan McKenzie had published a theoretical model suggesting that sedimentary basins were formed when the thick outer layer of the Earth, the lithosphere, was thinned by stretching. Matthews set out to test this model in the North Sea, a classic sedimentary basin and already an important oil province. First he had to organize the construction of an entirely new set of seabed seismic-recording systems. Then, after several expeditions testing the instruments and recording seismic data across the North Sea, the data were analysed using state-of-the-art computer synthetic seismogram methods, and integrated with geological observations of the subsidence history from fossils and sediments.

The results, published in 1983, showed for the first time that the crust beneath the North Sea was indeed thinned by stretching and that McKenzie’s model was a good description of the process. Typically for Matthews, he left his name off the authorship, giving the full credit to his graduate students. Almost immediately, this model was applied wholesale by both academia and the oil industry, and it continues to underpin our understanding

of sedimentary basins around the world and their important hydrocarbon potential.

In 1981, Matthews sensed that it was time to start something new because, as he put it, he began to bore his younger friends by saying gloomily when they came with bright suggestions that “we tried that 20 years ago, and it didn’t work”. So he turned his attention to the continental rather than the oceanic crust. His friend Jack Oliver at Cornell had shown that, with appropriate care, you could use conventional seismic-reflection techniques developed by the oil companies to look much deeper than they did — to the base of the crust and beyond. Matthews started a research group in Cambridge to do the same. Characteristically, rather than allowing this change of research direction to close down the Cambridge marine group that he had led so effectively, he chose to resign his university post to allow a marine-group successor to be appointed, and he put himself on soft money along with the rest of his new research group.

Thus was born the British Institutions Reflection Profiling Syndicate (BIRPS). It is a measure of his wry sense of humour that the second word was originally ‘universities’, but BURPS as an acronym was perhaps more than his research-council paymasters could stomach. The BIRPS group, under Matthews’ leadership, was as successful as his marine-group enterprise: it set new standards for deep crustal imaging, routinely recording complex and hitherto-unknown structure in the lower crust and upper mantle.

In 1990, after a warning heart attack the previous year, Matthews took early retirement and spent the rest of his life gently encouraging his many friends and former students. He continued his life-long enjoyment of sailing, first on a yacht and later, more genteelly, around the rural waterways of England on a small canal boat.

It is a measure of his mature leadership qualities and the confidence and scientific rigour that Matthews instilled in his students, that both of the research groups that he built — the marine group and BIRPS — have continued to thrive. And perhaps the strongest testaments to his scientific legacy are the many former students who are now in similar positions of leadership and responsibility elsewhere: between them they continue to make significant contributions to marine science around the world.

Robert S. White

Robert S. White is at the Bullard Laboratories, Madingley Road, Cambridge CB3 0EZ, UK.

Acoustic alarms reduce porpoise mortality

The most serious danger to dolphins and porpoises around the world is the threat from various forms of gill-net fishing. One potential way to reduce the number of deaths of marine mammals is the use of active acoustic alarms to warn animals about the presence of nets¹. Until now, acoustic alarms have not been tested in field experiments with sufficient statistical power². Here we describe a field experiment showing that acoustic alarms are effective at reducing the number of deaths of harbour porpoises (*Phocoena phocoena*) in sink gill-nets.

More than 80,000 small cetaceans are killed annually in coastal waters around the world, and at least two species (*Phocoena sinus* and *Lipotes vexillifer*) are in imminent danger of extinction because of fishing activities³. The number of documented cetacean population collapses due to fishing activities is probably underestimated because data are not available from most of the developing world, where extensive gill-netting is taking place.

Between 18 October and 15 December 1994, 15 commercial sink gill-net fishers from the coasts of New Hampshire and southern Maine took part in our experiment. Strings of 12 nets were tied together. Each net was roughly 92 m long and 4 m deep, and had a stretched mesh size of 15 cm. Whenever possible, the strings were soaked for 24 h and retrieved daily. Strings were set at least 100 m apart. Fishing took place on or near Jeffreys Ledge, New Hampshire, in an area closed to commercial sink gill-nets because of previous high incidental catches of harbour porpoises. Each vessel carried an independent observer, who changed vessels regularly throughout the experiment.

We used two visibly identical devices, some of which produced an acoustic alarm and others which were silent. Active devices were equipped with a switch activated by salt water that triggered the alarm on complete immersion so that neither fishermen nor observers knew whether strings had active or control alarms until the net was hauled. Each alarm was coded with a number that allowed us to track battery life, losses, malfunctions and the identity of alarms in the vicinity of porpoise by-catches. Thirteen alarms were placed 92 m apart, attached to the head rope at the end of each string and at each bridle, where individual nets were attached to each other. Each string had either a complete set of active alarms or a set of control alarms. The choice of active or control alarms for each string was made with a coin-toss the day before the string was retrieved and reset.

Observers carried a new set of dry alarms on board the vessel each day and replaced the alarms on strings of nets as they were retrieved. All alarms were changed on a string each time it was retrieved.

Active alarms emitted a broad-band signal with a fundamental frequency of 10 kHz, and a source level of 132 dB (reference pressure 1 μ Pa, measured at 1 m), well within the hearing range of harbour porpoises⁴ and harbour seals⁵. The signal lasted for around 300 ms and was repeated every 4 s. We chose the sound source levels so the alarms would be audible at 15 dB above ambient sound levels at 100 m (the length of one net) and to drop to ambient levels at 300 m. Ambient sound levels in the Jeffreys Ledge area range from 110 to 118 dB.

A total of 421 active and 423 control strings were set in similar locations and water depths with similar soak times. Two harbour porpoises were captured in active strings and 25 in control strings. In six control strings, two porpoises were caught in the same string; in all other cases only a single porpoise was taken. The maximum likelihood estimate of P (the probability of capturing at least one porpoise), was 0.025 for control and 0.0027 for active strings. These values were significantly different ($\chi^2 = 15.01$, $P = 0.0001$).

We captured similar quantities of two target species of the fishery (cod (*Gadus morhua*), $t = -0.43$, $P = 0.66$ and pollock (*Pollachius virens*), $t = 0.23$, $P = 0.82$) in control and active strings, as well as other commercial species. Of the non-target species that are important prey for the harbour porpoise, silver hake (*Merluccius bilinearis*) catches were similar in control and active strings ($t = -1.80$, $P = 0.08$), whereas Atlantic herring (*Clupea harengus*) were captured infrequently, but more often in control than active strings ($\chi^2 = 23.34$, $P = 0.01$). Seals caused damage to the fish catch with similar frequency in control and active strings. Two harbour seals (*Phoca vitulina*) were entangled in active strings and a single seal was caught in a control string.

At present we do not understand why the use of alarms produced such a dramatic reduction in porpoise catches. Studies of porpoises in a controlled setting suggest that alarms may alert porpoises to the presence of nets⁶, so that porpoises avoided the alarms and were therefore less likely to become entangled. An alternative hypothesis is that the effect was a result of the behaviour of herring, the primary prey of harbour porpoises in the Gulf of Maine⁷. Post mortem

examination showed that 17 of 19 porpoises had food remains in their stomachs. The species found most frequently were Atlantic herring (14 stomachs) and silver hake (10 stomachs). The presence of intact fish, flesh and bones, particularly from herring, indicated that many porpoises had been feeding just before entanglement. Herring were the only fish to show a significant difference in catch rate between active and control strings, with fewer herring taken in strings with active alarms. Clupeoid fishes have an unusual capacity for high-frequency hearing^{8,9} and it is possible that herring reacted to the alarms by avoiding the nets, thus reducing the number of porpoises becoming entangled while attempting to capture prey.

Our acoustic alarms reduced the incidental catch of harbour porpoises in sink gill-nets by an order of magnitude. The use of acoustic alarms would seem to be a promising method of reducing the number of harbour porpoises killed in sink gill-nets in the Gulf of Maine and offers hope for alleviating the bycatch problem for small cetaceans worldwide. Recent attempts to apply our results to local experimental fisheries have not been rigorously controlled, and some have produced mixed results, so the testing of alarms in other situations where odontocetes are threatened by gill-nets should proceed with careful experimental design and appropriate controls.

Scott D. Kraus

New England Aquarium,
Edgerton Research Laboratory, Central Wharf,
Boston, Massachusetts 02110, USA
e-mail: skraus@neaq.org

Andrew J. Read

Duke University Marine Laboratory,
Beaufort, North Carolina 28516, USA

Andrew Solow

Woods Hole Oceanographic Institution,
Woods Hole, Massachusetts 02543, USA

Ken Baldwin

Department of Ocean Engineering,
University of New Hampshire, Durham,
New Hampshire 03824, USA

Trevor Spradlin

Office of Protected Resources,
National Marine Fisheries Service,
Silver Spring, Maryland 20910, USA

Eric Anderson, John Williamson

New Hampshire Commercial Fishermen's
Association, Rye, New Hampshire 03780, USA
R. Barnaby, P. Ruell, K. Gestring and S. Drew facilitated data
collection aboard the fishing vessels.

- Jefferson, T. A. & Curry, B. *Biol. Conserv.* **67**, 167–183 (1994).
- Dawson, S. M. *Rep. Int. Whaling Comm.* **15**, 573–578 (1994).
- Perrin, W. F., Donovan, G. P. & Barlow, J. (eds) *Rep. Int. Whaling Comm. Spec. Iss.* **15** (1994).
- Andersen, S. *Invest. Cetacea* **2**, 256–259 (1970).
- Möhl, B. *J. Aud. Res.* **8**, 27–38 (1968).
- Kastelein, R. A., Goodson, A. D., Lien, J. & de Haan, D. in *Harbour Porpoises: Laboratory Studies to Reduce Bycatch* (eds Nachtigall, P. E., Lien, J., Au, W. W. L. & Read, A. J.) 157–167 (De Spil, Woerden, 1995).
- Recchia, C. R. & Read, A. J. *Can. J. Zool.* **67**, 2140–2146 (1989).
- Enger, P. S. *Comp. Biochem. Physiol.* **22**, 527–538 (1967).
- Popper, A. N. & Platt, C. *Nature* **280**, 832–833 (1979).

UK birds are laying eggs earlier

The evidence for global climate change and for its underlying anthropogenic causes is gathering rapidly. Over the past 11 years the active growing season of plants has advanced by roughly 8 days in northern latitudes¹. This evidence for increased photosynthetic activity is supported by the positive trend in the amplitude of the seasonal cycle in atmospheric CO₂ (ref. 2). The phenology of animal populations should also be affected by climate change, but to date there has been little evidence of this. Here we report that long-term trends in the seasonal distributions of laying dates of birds in the United Kingdom show a tendency towards earlier laying, consistent with the changes reported in growing season.

Since 1939, the nest record scheme of the British Trust for Ornithology has gathered more than a million records on the breeding performance of 225 species of birds in the United Kingdom. A network of 1,000 volunteer ornithologists provided details of nest site, habitat, contents at each visit and evidence for success or failure of each nest found³. Only a proportion of this information is available on computer databases and is sufficiently detailed to provide information on egg-laying dates.

We have now analysed 74,258 records from 65 species to investigate trends in the distributions of laying dates of the first egg in each clutch over the 25-year period from 1971 to 1995. We found significant trends towards earlier laying dates for 20 species (31%) with only one species laying significantly later. The shift towards earlier laying for the 20 species averaged 8.8 days, ranging from 4 to 17 days, s.d. = 3.4 (Fig. 1). The analyses of all species also show a tendency towards earlier laying (Fig. 2).

Species showing significant trends were

Figure 1 Temporal changes in laying dates for early-, mid- and late-season nesters. **a**, *Miliaria calandra*, $F_{1,23.5} = 5.30$, $P = 0.030$. **b**, *Phylloscopus collybita*, $F_{1,271} = 23.00$, $P = 0.0001$. **c**, *Pica pica*, $F_{1,38.9} = 61.89$, $P = 0.0001$. Laying date is numbered such that day 60 is 1 March, 121 is 1 May and so on. Points show annual means $\pm$ s.e.m. We analysed between 150 and 5,700 records per species. Laying dates were estimated with an accuracy of at least ± 5 days (ref. 3). For each species we selected the most significant of three mixed linear models with Year (for example, **a** and **b**), Year² (**c**), or Year and Year² combined, fitted as continuous fixed effects; Year was also fitted as a categorical random effect to contend with the non-independence of observations within years.

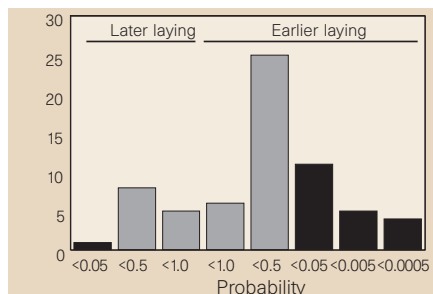
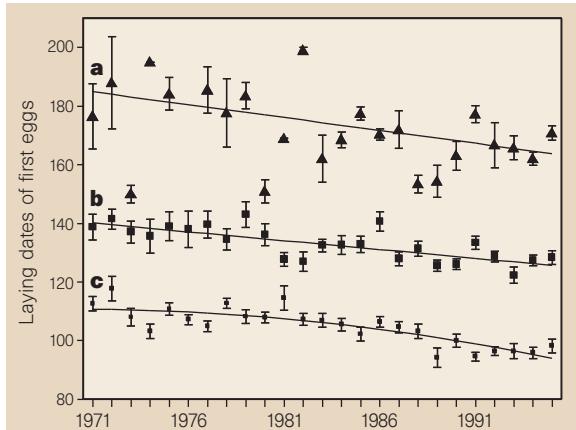


Figure 2 Frequency distribution of P values of temporal trends in laying dates for 65 species of UK birds from 1971 to 1995. Significant trends are shown in black, non-significant in grey. Trends towards earlier laying were found in 51 species and towards later laying in 14 species, a significant difference at $P = 0.000005$ (binomial test).

not confined to any one ecological or taxonomic type, and comprised water birds (*Haematopus ostralegus*, *Numenius arquata*, *Tringa totanus*), resident insectivores (*Cinclus cinclus*, *Troglodytes troglodytes*, *Aegithalos caudatus*, *Sitta europaea*, *Sturnus vulgaris*), migrant insectivores (*Anthus trivialis*, *Phoenicurus phoenicurus*, *Sylvia communis*, *S. atricapilla*, *Phylloscopus sibilatrix*, *P. collybita*, *P. trochilus*), corvids (*Pica pica*, *Corvus corone*) and seed-eaters (*Fringilla coelebs*, *Carduelis chloris*, *Miliaria calandra*).

The species also covered a wide range of nesting times, from early to late season (Fig. 1), which suggests that the recorded trends were not due to changes in the behaviour of observers. The only significant trend towards later laying was for the stock dove (*Columba oenas*), which nests opportunistically throughout the year and may therefore be a special case.

Trends towards earlier laying times are expected if ambient temperatures rise earlier in the year. There are fitness benefits to nesting early, and food availability is often the important determinant of laying date⁴. Average flowering and leafing dates may be advanced by high spring temperatures^{5,6} and these are likely to have pronounced

effects on the availability of the arthropod food supplies for birds. There is evidence that two species of waders in The Netherlands nest earlier in warmer springs⁷. Amphibians in Britain have also been shown to spawn earlier in recent years as spring temperatures have risen⁸.

The tendency to nest or reproduce earlier may be a more general phenomenon for wildlife in Britain. This could have considerable consequences for their ecology and conservation. For birds, earlier nesting could be beneficial if juvenile survival is enhanced by a prolonged period before winter. Conversely, birds may be adversely affected if they become unsynchronized with the phenology of their food supplies.

Humphrey Q. P. Crick
Caroline Dudley, David E. Glue
David L. Thomson

British Trust for Ornithology,
The National Centre for Ornithology,
Thetford, Norfolk IP24 2PU, UK

1. Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G. & Nemani, R. R. *Nature* **386**, 698–702 (1997).
2. Keeling, C. D., Chin, J. F. S. & Whorf, T. P. *Nature* **382**, 146–149 (1996).
3. Crick, H. Q. P. & Baillie, S. R. A Review of the BTO's Nest Record Scheme (Brit. Trust Ornithol., Thetford, 1996).
4. Perrins, C. M. *Acta XIX Congr. Int. Ornithol.* **1**, 892–899 (1988).
5. Sparks, T. H. & Carey, P. D. *J. Ecol.* **83**, 321–329 (1995).
6. Fitter, A. H., Fitter, R. S. R., Harris, I. T. B. & Williamson, M. H. *Func. Ecol.* **9**, 55–60 (1995).
7. Kruk, M., Noordervliet, M. A. W. & ter Keurs, W. J. *Biol. Conserv.* **77**, 213–218 (1996).
8. Beebe, T. J. C. *Nature* **374**, 219–220 (1995).

CO₂ increases oceanic primary production

The regulation of oceanic primary production of biomass is important in the global carbon cycle because it constitutes 40% of total primary production on Earth¹. Here we present results from short-term experiments in the nutrient-poor central Atlantic Ocean. We find a small but significant stimulation of primary production (15–19%) in response to elevated CO₂ concentrations that simulate the CO₂ rise in surface waters that will occur over the next 100–200 years.

Traditionally, inorganic carbon has not been regarded as a limiting factor for oceanic production^{1–3}. Most inorganic carbon is present as HCO₃[–] (~2 mM) whereas free CO₂ comprises only ~10 μ M at atmospheric equilibrium (25 °C). The main carboxylating enzyme in marine phytoplankton requires 25–35 μ M CO₂ to saturate², so phytoplankton species that cannot concentrate CO₂ internally^{2,4} by active transmembrane transport may face rate limitation of photosynthesis by CO₂ diffusion to the cell surface. Even algae capable of using HCO₃[–] as an external carbon source can be limited by low availability of zinc, owing to its catalytic role

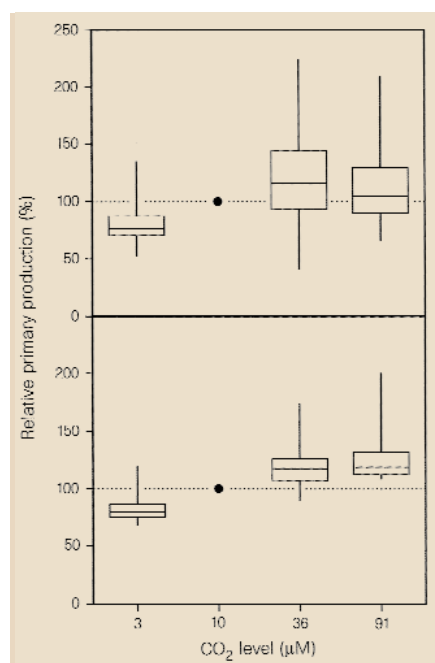


Figure 1 Oceanic primary production at reduced and increased CO₂ concentration (percentage of production at ambient CO₂). Measurements are triplicates at 18 stations along a transect in the Atlantic Ocean (35° S 49° W to 24° N 29° W) on board the Spanish BIO Hesperides. Upper panel, surface layer; lower panel, deeper layer of maximum chlorophyll content. CO₂ concentrations were: ambient 10 ± 0.3 μM, intermediate 36 ± 1.2 μM, high 91 ± 3.1 μM and low 3 ± 0.1 μM. Horizontal lines, median values; boxes, 25 and 75 percentiles; vertical lines include 10–90 percentiles. The coefficient of variation of triplicates for production rates averaged 11% in surface and 7% in deep layers. Primary production rates at elevated intermediate CO₂ concentration (36 μM) were significantly higher than those at ambient levels for surface and deeper layers ($P=0.028$ and $P<0.001$, respectively; one-tailed signed rank test).

in the enzyme carbonic anhydrase⁵. Culture experiments and diffusion models with marine phytoplankton have therefore suggested that photosynthesis and growth can be limited by the supply rate of CO₂ (ref. 4). But culture experiments with a few species cannot provide a fair description of the natural response of the mixed phytoplankton assemblage in the ocean where many species are involved and cell density, nutrient status and physiological acclimation may differ widely from those in culture.

We tested whether primary production was limited by ambient CO₂ concentration at 18 stations along a transect in the Atlantic Ocean. Primary production was measured as short-term ¹⁴C-incorporation into organic matter in samples collected from the surface waters (~5 m) and from the deeper layer (30–150 m) with maximum chlorophyll content. We illuminated surface samples at 350 μmol photon m⁻² s⁻¹ and deeper samples at 50 μmol m⁻² s⁻¹ for 2 h in an incubator kept at ambient temperature.

We elevated or reduced the CO₂ concentration relative to the ambient level by slightly changing the pH, adding HCl or NaOH. This procedure is straightforward and can change CO₂ concentrations without altering the total inorganic carbon pool. The decline in pH along with the CO₂ rise may influence primary productivity, but this pH decline would also occur naturally with the future CO₂ rise. The alteration of pH may also influence trace metal bioavailability⁵, but the duration of the experiments was so short (2 h) that they are unlikely to influence ¹⁴C incorporation.

Overall, primary production at the 18 sites responded significantly to manipulation of the CO₂ concentration (Fig. 1). In surface waters, median primary production at low CO₂ (3 μM) was 75% of the level at ambient CO₂ (10 μM) and median primary production at elevated CO₂ (36 μM) was 115% of the ambient level. Likewise, in deeper chlorophyll-rich layers, median primary production was 78% at low CO₂ and 119% at elevated CO₂ as compared with ambient primary production. We saw no further increase in primary production at the highest CO₂ level tested (91 μM).

The response to CO₂ manipulation varied substantially among different sites, as expected considering that the composition of the phytoplankton communities was likely to vary along the transect. At some sites primary production did not respond to CO₂ elevation whereas at other sites primary production more than doubled. Rates of primary production at ambient CO₂ in the surface (0.02–0.44 mg C m⁻³ h⁻¹) and deeper layers (0.05–0.86 mg C m⁻³ h⁻¹) also varied considerably. There was no systematic relationship between the response of primary production to CO₂ manipulation and the geographical location or magnitude of primary productivity.

The CO₂ concentration elevation to 36 μM corresponds roughly to the level expected in the surface ocean at atmospheric equilibrium (855 p.p.m., 25 °C) in year 2100, according to the IPCC 'business as usual' model⁶. Our results indicate that the median primary production could increase by about 15–19% in response to the CO₂ rise, with other factors remaining constant. The variable response to CO₂ elevation along the transect also suggests that the stimulation could be markedly higher (perhaps double) in some instances, probably dominated by phytoplankton species without CO₂-concentrating mechanisms. However, the long-term growth response to CO₂ elevation is usually much lower than the short-term photosynthetic response determined here. This has been shown frequently for terrestrial plants³.

We propose that the overall response of oceanic primary production to the CO₂ rise would be relatively small, whereas the influ-

ence on the species composition of the phytoplankton assemblages could be profound, depending on the kinetics of carbon use.

Mette Hein, Kaj Sand-Jensen

Freshwater Biological Laboratory,

University of Copenhagen,

Helsingørsgade 51, DK-3400 Hillerød, Denmark

e-mail: mette_hein@dk-online.dk

1. Falkowski, P. G. *Photosynth. Res.* **39**, 235–258 (1994).
2. Raven, J. A. & Johnston, A. M. *Limnol. Oceanogr.* **36**, 1701–1714 (1991).
3. Bowes, G. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **44**, 309–332 (1993).
4. Riebesell, U., Wolf-Gladrow, D. A. & Smetacek, V. *Nature* **361**, 249–251 (1993).
5. Morel, F. M. M. *et al. Nature* **369**, 740–742 (1994).
6. Houghton, J. T., Jenkins, G. J. & Ephraïm, J. J. (eds) *The IPCC Scientific Assessment on Climate Change* (Cambridge Univ. Press, 1990).

Structural biology and phylogenetic estimation

When reconstructing evolutionary trees from DNA sequences, it is often assumed that increasing the amount of sequence will improve the phylogenetic estimate^{1,2}. This is based on the notion that historical 'signal' will rise above misleading 'noise' as more sequence is gathered. Our analysis of mitochondrial genomes fails to support this assumption, but suggests a way to select objectively for data with maximum 'signal-to-noise' potential.

We carried out a phylogenetic parsimony analysis³ of the entire protein-coding portion (12,234 base pairs) of the mitochondrial genome of 19 taxa whose interrelationships are widely accepted (Fig. 1a). Not only did we fail to obtain the expected phylogenetic tree from this large data set, but found compelling bootstrap support for the incorrect placements (Fig. 1b). This suggests that misleading signals will not always disappear as more data are collected.

When a subset of sites associated with residues important for protein folding was subjected to phylogenetic analysis, the expected tree resulted with strong bootstrap support. Phylogenetic analysis of DNA sequences might therefore be improved by incorporating structural and functional considerations into inference models^{4–6}.

To determine which of our 12,234 sites were responsible for the misleading signal, we superimposed the sequence data onto the accepted tree and measured, using the retention index^{7,8}, how well each site fitted that tree. We found that among codons, nucleotides at third positions produced the poorest fits; among genes, *NADH2* produced the poorest fit; and isoleucine, leucine and valine produced the poorest fits of the amino acids. Sites coding for residues with aliphatic side chains produced poorer fits than did those coding for any other type

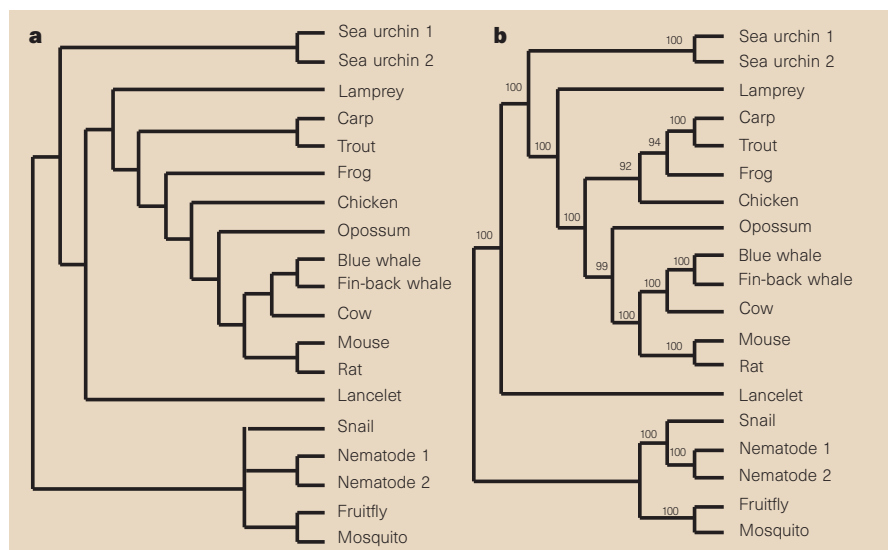


Figure 1 a, The widely accepted pattern of relationships for the 19 taxa analysed. Species names: mouse, *Mus musculus*; rat, *Rattus norvegicus*; cow, *Bos taurus*; fin-back whale, *Balaenopterus physalus*; blue whale, *Balaenopterus musculus*; opossum, *Didelphis virginiana*; chicken, *Gallus gallus*; frog, *Xenopus laevis*; carp, *Cyprinus carpio*; trout, *Oncorhynchus mykiss*; lamprey, *Petromyzon marinus*; lancelet, *Branchiostoma floridae*; sea urchin 1, *Paracentrotus lividus*; sea urchin 2, *Strongylocentrotus purpuratus*; fruitfly, *Drosophila yakuba*; snail, *Cepaea nemoralis*; mosquito, *Anopheles gambiae*; nematode 1, *Ascaris suum*; nematode 2, *Caenorhabditis elegans*. **b**, Bootstrap consensus tree based on a parsimony analysis of 12,234 nucleotide sites coding for the 13 mitochondrial genes. Numbers show percentage bootstrap support at each node.

of side chain; sites coding for charged residues produced better fits than did those for uncharged residues; and sites coding for hydrophilic residues produced better fits than did those coding for hydrophobic ones (Fig. 2).

On the basis of these analyses, we were able to identify functional classes of sites that were phylogenetically reliable. We found that parsimony analysis of a subset of sites comprising first and second codon positions, modally coding for proline, cysteine, methionine, glutamine and asparagine, yielded the expected tree with strong

bootstrap support for most nodes. Although this cannot serve as an independent test (the expected tree was used to discover sites that were resilient), it is noteworthy that most of the resilient sites are associated with amino-acid residues important for tertiary protein structure.

Tree-based explorations of sequence variation such as this show considerable promise for both evolutionary and structural biology. The distributions and incidences of change implied by an accepted set of evolutionary relationships can be physically mapped onto a structural model of a protein or RNA

product^{9,10} to shed light on the constraints and opportunities that have characterized the evolution of the underlying sequence. This information can be used both to refine phylogenetic models and to provide structural biologists with clues about the relative importance of particular co-varying combinations of residues for protein structure, function and folding.

Comparative morphologists have long acknowledged that character co-variation among organisms is affected not only by a shared history, but also by functional requirements. To work as an integrated unit, the sub-components of a complex morphological character must necessarily co-vary. There is every reason to believe that molecules are similar. As the database of protein structures continues to grow, and as sequences from an ever broader taxonomic range of organisms become available for a number of proteins, we can look forward to a better understanding of the historical and functional constraints that act on macromolecules and, as a consequence, to more realistic biochemically based models of change from which to infer evolutionary trees.

Gavin J. P. Naylor

Department of Zoology and Genetics,
Iowa State University, Ames, Iowa 50011, USA
e-mail: gsnaylor@iastate.edu

Wesley M. Brown

Department of Biology, University of Michigan,
Ann Arbor, Michigan 48109-1048, USA

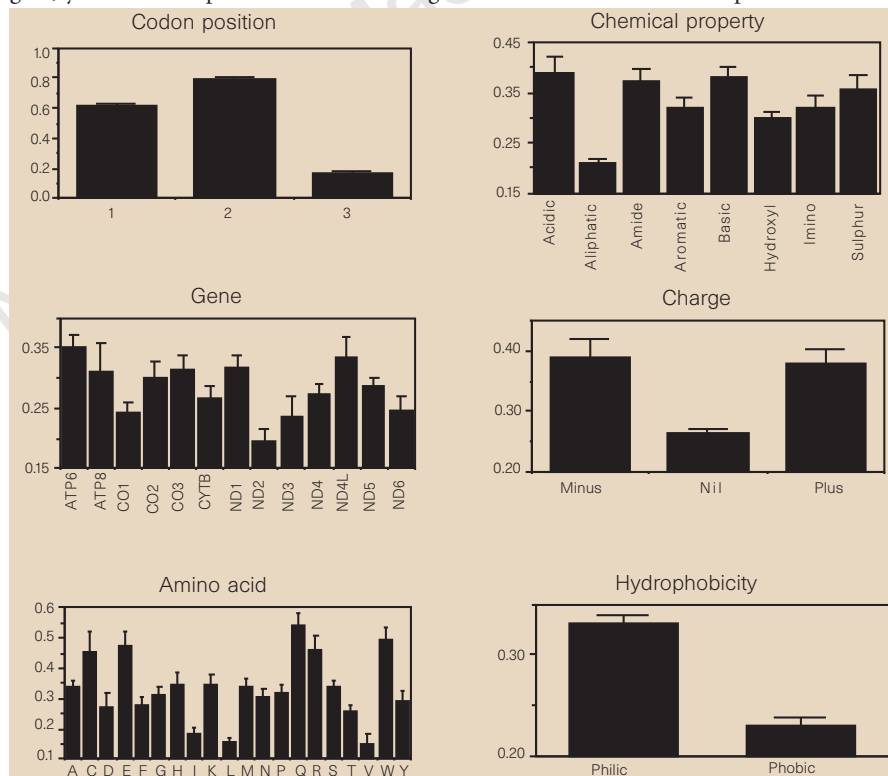


Figure 2 Relationship between functional characteristics and phylogenetic fit of a site for the accepted tree. Fit was assessed using the retention index²⁸. The relative effects of the different levels of each influence are plotted against log retention index (ordinate) as response sample means. Bars show s.e.m. Gene abbreviations: ATP6, ATP8, ATPase subunits 6 and 8; CO1-3, cytochrome oxidase subunits 1-3; CYTB, cytochrome B; ND1-6 and ND4L, NADH dehydrogenase subunits 1-6 and 4L. Amino acids are listed as single-letter code.

- Churchill, G. A., von Haessler, A. & Navidi, W. C. *Mol. Biol. Evol.* **9**, 753-769 (1992).
- Huelsenbeck, J. P. & Hillis, D. M. *Syst. Biol.* **42**, 247-264 (1993).
- Swofford, D. L. *PAUP 3.1.1* (Smithsonian Inst., Washington DC, 1993).
- Irwin, D. M., Kocher, T. D. & Wilson, A. C. *J. Mol. Evol.* **32**, 128-144 (1991).
- Adkins, R. M., Honeycutt, R. L. & Disotell, T. R. *Mol. Biol. Evol.* **13**, 1393-1404 (1996).
- Perutz, M. F. *Mol. Biol. Evol.* **1**, 1-28 (1983).
- Archie, J. W. *Syst. Zool.* **38**, 253-269 (1989).
- Farris, J. S. *Cladistics* **5**, 417-419 (1989).
- Goulding, B. in *Non-neutral Evolution* (ed. Goulding, B.) 126-139 (Chapman and Hall, New York, 1994).
- Schnare, M. N., Dammerger, S. H., Gray, M. W. & Gutell, R. R. *J. Mol. Biol.* **256**, 701-719 (1996).

The planetary piggy bank

Nature's Services: Societal Dependence on Natural Ecosystems

edited by Gretchen C. Daily

Island: 1997. Pp 392. \$49.95 (hbk); \$24.95, £19.95 (pbk)

The Work of Nature: How the Diversity of Life Sustains Us

by Yvonne Baskin

Island: 1997. Pp 263. \$25, £19.95

Mary Power

Within the past few months, these two new books and an article in *Nature* have addressed the value of ecosystem services. What does nature do for the human enterprise, and how much are its services worth, in cash? Posing this question may initially seem crass or misguided, recalling the greed and foolishness of King Midas, whose myth reverberates today in the anecdote closing a chapter by Rosalind Naylor and Paul Ehrlich in *Nature's Services*.

After a speaker at a US Federal Trade Commission meeting dismissed the threat of global warming by pointing out that agriculture and forestry account for only 3 per cent of US gross national product, the audience overheard a distinguished economist comment: "What does this genius think we're going to eat?"

Practical, ethical and aesthetic problems of putting a price on nature, as well as the ultimately incalculable value of natural life-support on our planet, are acknowledged upfront in these two books, as well as by Robert Costanza and colleagues in their *Nature* article (387, 253; 1997 (see also News and Views, 387, 231; 1997)).

Nevertheless, all these authors make compelling cases for why the valuation exercise is crucial. Ecosystem services, defined in *Nature's Services* as "the goods and life support functions provided by natural ecosystems and their species that sustain and fulfill human life", have long been undervalued and understudied. By attempting to price services such as water purification, pest control or climate amelioration, these authors provide some of the first quantitative estimates — at ecosystem, regional and global scales — of "what is lost when we degrade the free services that nature provides". Tailored to specific regions or watersheds, such information would greatly improve our ability to take explicit account of the tradeoffs we make in the course of decisions about land use.

The two books and the *Nature* article are lucid, substantive and authoritative. Taken together, there is surprisingly little redundancy, with different and complementary emphases.

Costanza gathered economic and ecological colleagues to develop a total valuation

for the world's ecosystem services and natural capital. Their estimate, \$33 million million a year, is the sum of the estimated annual global value of 17 types of services distributed among 21 biomes. This is almost certainly an underestimate, given the lack of information on some biomes or types of services, several conservative economic assumptions and, most importantly, the static nature of this snapshot estimate. As supplies diminish, values of commodities or services increase, exponentially to infinity if there is no substitute.

The Costanza estimate is clearly not robust, given its sensitivity to complex and dynamic ecological and economic forces. Yet it is extremely useful, if only to cause reflection that this conservative estimate is already twice the global annual gross national product.

In *Nature's Services* the emphasis shifts from a total valuation to estimating the marginal values of ecosystem services. Marginal values are defined in the book as "prices people are willing to pay for each successive increase in an ecosystem service, or the value consumers place on the last unit purchased".

This estimate can be broken into two parts: the increment (or decrement) in human welfare gained from a small change in an ecosystem service, and the change in the ecosystem service caused by a small change in policy. These values are chal-

lenging to measure and clearly depend on where slopes are measured along curves relating the variables. Therefore, like total valuation estimates, marginal values also depend on context and assumptions.

Nevertheless, the task of measuring the marginal value of a particular ecosystem service in a particular region is extremely heuristic. The approach focuses the problem so that quantitative models with operationally defined terms and explicit testable assumptions can be applied to evaluate land use issues and eventually (dare we hope?) assist in predicting the ecological-economic system's behaviour under various scenarios. The challenge of determining the way in which small changes in policy affect ecosystem states and processes, and how these affect human society, sets an important research agenda.

Nature's Services is an important first step in this line of enquiry. It moves the study of ecosystem services beyond its earlier phase, during which a handful of workers presented global case studies of ecosystem services and pleaded for recognition of their importance.

Nature's Services begins with Gretchen C. Daily's overview of ecosystem services, in which she frames and defines key issues. Philosophical and historical comments follow by Daily, N. Myers, J. S. Reichert, J. P. Myers, H. A. Mooney and P. R. Ehrlich. Two



Life in the freezer

Glaciologist Charles Swithinbank worked for many years for the US Antarctic Program. His career spanned a period of unprecedented change: "Schooled, as I was, in the British tradition of traveling the hard way, with man-hauled sledges, dog teams, and spartan rations, my move to the US plunged me into a world of giant icebreakers [above], aircraft, tractors,

gourmet food, and a US Naval Task Force struggling to come to terms with the eccentric ways of scientists." Swithinbank describes his "adventures in the pursuit of science" in *An Alien in Antarctica: Reflections upon Forty Years of Exploration and Research on the Frozen Continent* (McDonald and Woodward, \$49.95).

chapters on economics follow in which L. H. Goulder and D. Kennedy review alternative approaches to valuation and R. Costanza and C. Folke discuss how valuations are complicated by context, including the many different goals of societies.

The next section contains five excellent overviews of feedback between life and climate (S. E. Alexander, S. H. Schneider and K. Lagerquist); biodiversity and ecosystem function (D. Tilman); and the services provided by soils (G. C. Daily, P. A. Matson and P. M. Vitousek), pollinators (G. P. Nabhan and S. L. Buchmann) and natural enemies contributing to the control of agricultural pests (R. L. Naylor and P. R. Ehrlich).

All these chapters succeed in orienting general readers to the topic, yet also provide comprehensive up-to-date summaries of recent research and guides to the relevant literature. The authors use sometimes ingenious methods of estimating absolute and marginal values of particular ecosystem services at regional and global scales.

The third section reviews services supplied by important biomes: marine (C. H. Peterson and J. Lubchenco); fresh water (S. Postel and S. Carpenter); forest (N. Myers); and grassland (O. E. Sala and J. M. Paruelo).

Postel and Carpenter point out that while fresh water has no substitute, its value under certain limiting conditions can be estimated from the replacement costs of using desalinated sea water. This practice is currently largely confined to the Persian Gulf where nations convert oil wealth to water at a cost of \$1–\$2 per cubic metre. In the less water-starved and less wealthy arid western United States, their analysis shows that today's marginal value of water for recreation, fisheries and other in-stream uses often exceeds the marginal value of water diverted for agriculture, so that "if the water market were freer, fewer diversions would be economically justified".

The next section contains miscellaneous case histories including a chapter on marine resource extraction (L. Kaufman and P. Dayton); a survey of the value of global biodiversity (N. Myers); tradeoffs between environmental conservation and development in a community in the Rocky Mountains in the United States (A. Wilcox and J. Harte); a discussion of the uses of natural and reconstructed wetlands (K. C. Ewel); and an analysis by R. M. Cowling, R. Costanza and S. I. Higgins of whether beating back the invasive exotic vegetation threatening the fynbos (a hard-leaved, fire-prone shrubland) ecosystem of South Africa can be economically justified (it can).

A particularly fascinating chapter in this section by K. S. Bawa and M. Gadgil reviews ecosystem services used by people in subsistence economies, and discusses the feasibility of enlisting these people in efforts to preserve their local biodiversity. Follow-

ing R. F. Dasmann, Bawa and Gadgil distinguish "ecosystem people" who rely on local ecosystems for most of their resources from "biosphere people", such as urbanites whose resources are drawn from much of the planet. Environmental degradation caused by resource extraction is generally invisible to biosphere people, but has immediate and sometimes tragic consequences for ecosystem people, such as the ecological refugees displaced by dams, deforestation or anthropogenic desertification.

Bawa and Gadgil point out that ecosystem people have conserved biodiversity and natural resources "when these serve the long-term interests of a small, well-knit human group". They offer practical ideas for fostering social and economic conditions that would permit these people to remain stewards and beneficiaries of their lands.

Yvonne Baskin's *The Work of Nature* is a well-written popular account of the recent investigations of community ecologists and ecosystem scientists whose findings are relevant to ecosystem services. With many vivid examples, Baskin reviews research that illustrates how species interactions influence ecosystems, emphasizing that some species that appear unnecessary or 'redundant' today may prove crucial tomorrow as conditions or circumstances change.

Her chapters reviewing keystone species and community interactions are gems, as is her lucid treatment of the adverse impact of invading species on the health of forests, fisheries or humans. Later she discusses ecosystem-level investigations of how soil and vegetation interact with climate and hydrology. She then reviews the fascinating studies of N. Owen-Smith, P. Martin, and S. Zimov, T. and M. Chapin and their colleagues who, in Africa, North America and Beringia respectively, have built cases for continental scale impacts of prehistoric humans who eliminated megaherbivores.

In her chapter "Do we still need Nature?" Baskin discusses the failure of the \$150 million Biosphere II project, in which the atmosphere suffered from overproduction of carbon dioxide and underproduction of oxygen. She relates several other cautionary tales of failed engineering fixes that attempted to replace free natural services lost as ecosystems degraded.

With specific case studies such as the pollution of estuaries by cage culture of salmon, she illustrates problems that arise when local owners and beneficiaries are not paying for the loss of services from ecosystems that they overwhelm and degrade. As Daniel Schneider has pointed out in his work on the conversion of the Illinois River floodplain to cornfields, in contrast to Hardin's "tragedy of the commons" model, the most severe environmental degradation can follow the privatization of formerly public lands.

Both books are 'must reads' for teachers,

students, scientists and citizens at all levels of expertise. The valuation of nature should be considered and widely debated by all concerned with humanity's future.

Pricing nature may disturb and depress 'green' readers. They should keep reading, for these authors are not among those "for whom every value can be made monetary", to use Naylor and Ehrlich's phrase. The authors do not believe that nature's only purpose is to comfort, support or amuse humans. They have simply confronted the reality that, in the world as it is today, putting a minimum monetary estimate on nature's services is the only way to slow, and hopefully eventually to stop, the loss of its crucial life-support functions. □

Mary Power is in the Department of Integrative Biology, University of California at Berkeley, Berkeley, California 94720, USA.

Science in the service of the Raj

The Science of Empire: Scientific Knowledge, Civilization, and Colonial Rule in India

by Zaheer Baber

State University of New York Press: 1996.

Pp. 298. \$71.50, £55.75 (hbk); \$23.95, £18.75 (pbk)

Mapping an Empire: The Geographical Construction of British India, 1765–1843

by Matthew H. Edney

University of Chicago Press: 1997. Pp. 436.

\$35, £27.95

Ehsan Masood

Step onto any main thoroughfare in New Delhi, Mumbai (Bombay), Lahore or Karachi, and you'll find that the traffic drives on the left, just as it does in Britain. Step into an office, and you'll be offered tea with milk and sugar served by uniformed men still quaintly referred to as 'peons', the name given to the native servants of the Raj.

The imperial crown fell with the creation of India and Pakistan on 14 and 15 August 1947. But, by most accounts, the shadow of the Union Jack, the gaberdine suit and the khaki helmet still looms large.

English remains the language of business and education. Application forms are completed 'in triplicate'. 'Oxford' is a name used by garment manufacturers and nursery schools. The school leaving examination is still called a 'matriculation'. A university student with high enough grades will graduate with 'honours'. And, should he want to become a scientist, he may end up working for the Council for Scientific and Industrial Research.

Two books on the role of science in the spread of empire published in time for the fiftieth anniversary of independence confirm

that the stamp of the crown is too well defined to be erased by the ritual renaming of streets and towns, the compulsory wearing of non-Western attire, or crude attempts to abolish English as the main language of education.

The books by Zaheer Baber and Matthew H. Edney are part of the already substantial literature on empire to emerge from the colonial archives at, among others, the British Library Oriental and India Office Collections in London.

Some historians do not consider sifting through such records as 'real' historical research. They argue that such an exercise is not so much research as stamp-collecting, as the material is already in one place, and properly indexed. They also believe that such investigations reflect only part of a story as they rely almost exclusively on 'secondary' source materials; indeed, from a single source — the colonial power. They also argue that the conclusions drawn from such studies are hardly new, as each survey repeats previous findings that colonial powers exploited those in their control for economic and territorial gains.

There is merit in such an argument. But it would be unfair to dismiss these books as unwanted additions to the overcrowded market for 'secondary-level' studies on the British Empire. For that is not their aim.

The books' novelty lies not in their conclusions on the exploitative nature of empire. Rather, they provide a glimpse of the fruits that lie ahead if historians of science, and sociologists of science, start to acknowledge each others' existence. The virtual absence of people in the field of science studies from last month's international congress of historians of science in Liège, Belgium, suggests, for example, that such an accommodation has some way to go.

Why should this be the case? Until now, many historians of science have been content to piece together specific examples from research in the past, without necessarily recounting the prevailing social, cultural, religious and political forces that may have shaped the nature of that science. This irritates sociologists of science who argue that digging out the minutiae of science in India, China, the Middle East or the Ottoman Empire is meaningless unless understood in the prevailing social context. But sociologists, too, have so far shown little vision in this area, and seem to prefer to go into combat with natural scientists about questions on the nature of science.

It is, therefore, heartening to see Baber, a sociologist at the National University of Singapore, break free from the 'Science Wars' by attempting to analyse the politics of science in colonial India in *The Science of Empire*. In the second book, *Mapping an Empire*, Edney, associate professor of geography and anthropology at the University of Southern Maine, performs a similar exercise on one specific area of science in British India: cartography.

What Baber and Edney have done is effec-



Government House, Calcutta, built in 1798–1803 for Governor-General Lord Wellesley, whose recall was partly due to the extravagant outlay.

tively to reconstruct the history, debates and politics of science in colonial India. The strengths of both studies are the detailed pictures that emerge from the background to the setting up of major institutions of science.

Baber and Edney demonstrate that science in colonial India was no abstract pursuit for gentlemen of leisure that happened by accident. Rather it was a multi-purpose tool with clear applications in mind. These included providing scientific advice to government, assisting in the process of wealth creation, and re-educating the 'natives' in the virtues of science and reason, in place of their own 'irrational' beliefs and practices.

It is fascinating to read how many of the debates described are still as relevant today as they were two centuries ago.

While Edney concentrates on institutions concerned with cartography, Baber takes a broader look at, among others, the Royal Asiatic Society — modelled on the Royal Society of London — the Royal Botanic Garden of Calcutta, and the Board of Scientific Advice, an expert advisory committee that coordinated research across India. The authors show that each institution emerged for a specific reason, which invariably revolved around the idea of easing the process of colonial administration, rather than merely pursuing good science.

India officially became a British colony in 1857. But it was a colony in all but name up to a century earlier while being administered by the powerful East India Company. The company had first arrived in the early 1600s as a trading company. It quickly prospered, and became involved in local politics. In 1757 it lent its large private army to a rebellion against the ruler of Bengal, and found itself in control of India's most prosperous province. Further battles followed, and the company — a confusing name for what was effectively a powerful military force — acquired more territories. The British government, not oblivious to these developments, increased its control over the company's affairs in 1813, before taking over in 1857.

The growth of the company's influence

across India generated the need for accurate geographical information. And one of the earliest applications of science in India was in attempts to survey the vast new territories falling under the company's jurisdiction.

Edney describes in detail how the mapping of India began in 1765 after the fall of Bengal, and how the surveys were supported by the East India Company as a means of establishing the boundaries of empire. "To govern territories," says Edney, "one must know them."

But the company was always a reluctant investor in research, particularly research that had no clear application. And, despite the rhetoric about introducing 'scientific method' to India, Edney reveals that the surveying was a disorganized affair that lacked a consistent method until as late as the 1830s, and was completed only after a series of technological fixes. Baber confirms the company's early hesitancy in funding science and shows that officials drew on the advice of research scientists only when problems in governance began to mount.

A series of devastating famines in Bengal in the second half of the eighteenth century was one such problem in which science was asked to play a major role. The famines began just after 1757, and led to massive crop failures, starvation and flight from the land. Faced with the prospect of a financial crisis from falling agricultural revenues, the company's directors welcomed a proposal in 1787 from an army officer and amateur botanist to set up the Royal Botanic Garden of Calcutta, not as a centre for classifying rare plants, but as a money-spinning venture.

The idea was cleverly sold to the directors of the East India Company as an experimental station to cultivate commodities with high revenue potential, such as spices, herbs, dyes, drugs, coffee and tea, and to replace revenues lost from failed crops, such as rice and cotton. This can with hindsight be considered an early example of technology foresight and, within four years of the setting up of the Calcutta garden, similar centres had sprung up in Madras and Bombay.

Baber points out that the network of gar-

dens was not set up to tackle the famines. These would continue for more than a century, killing hundreds of thousands, and turning one-third of Bengal into a wasteland, before research was brought to bear on the problem.

Science returned to the scene of the famines with the setting up in 1880 of a 'scientific commission' of experts. The commission concluded that the problem would be better understood if each province contained a dedicated agricultural research department, and India's existing network of agricultural research centres was born. But Baber says that the scientists did not have it all their own way. Recommendations that could harm Britain's interests, such as one to reduce dependency on agriculture by developing industry in India, were ignored.

Meanwhile, the famines continued, wiping out all livestock and leaving 6 million people destitute in 1900. A new viceroy, George Curzon, was more determined than ever to apply science and technology to alleviating famine. Agricultural research was strengthened further. Pay and conditions for scientists were improved. And in 1902, Curzon took the additional step of setting up a Board of Scientific Advice to advise the government and coordinate research relating to the economy and agriculture.

This board, says Baber, "bore a striking resemblance to the 'Scientific Advisory Council' proposed for Britain three decades earlier by Alexander Strange, and Norman Lockyer". Lockyer, the first editor of *Nature*, was excited about the project, and covered the founding of the board in the journal (67, 568; 1903), as did some newspapers.

But the board did little to alleviate the famines, which appear to have ceased a few years before the end of empire. Baber believes that the colonial administrators knew full well that the famines had begun more than a century earlier as a result of an onerous tax burden, coupled to the neglect of irrigation systems. Science, deliberately or otherwise, proved a useful diversionary tactic.

In addition, the board's usefulness as an analogue for a similar committee in Britain remains debatable, given its inappropriately bureaucratic nature. The board's annual reports were communicated through the Secretary of State for India to the Royal Society. The society would then consult another advisory committee set up to liaise between the administrations in Britain and India. This committee would report back to the colonial office in India, providing any appropriate advice.

Both books contain many other examples of how institutions of science emerged during the colonial period in response to the political and economic problems of the day. Baber also attempts to provide information on science in ancient and mediaeval India. But, lacking access to good primary source material, his coverage contains little of the

social and political context that so successfully illuminates his later chapters.

This, in a sense, brings us back to where we started. An understanding of the social context of science in different societies requires a mastery of history, science and language. The historiography of science in European colonial times is relatively easy to research, as most of the source material is well documented. An additional bonus for researchers from Europe and the United States is that it is in English. The development of science in non-Western empires, though, is more difficult to research.

Step back to before the sixteenth century, and the picture becomes hazy, as source material has yet to be properly catalogued. Some sociologists increasingly appear to want to pursue this research. But they lack the prerequisite training in language and history. On the other hand, the handful of those who have such skills—the professional historians of science—consider the pursuit of context to be a diversion from their job as historians. A meeting of minds is on few agendas.

This is unfortunate, as concerted cohabitation may one day lead to an assault on the one question that both groups are keen to answer: why did the scientific revolution happen in the West? Until this marriage takes place, we may never know. □

Ehsan Masood is a science writer on *Nature* and is writing a book on the politics of science in Pakistan.

A private function

Why is Sex Fun? The Evolution of Human Sexuality

by Jared Diamond

BasicBooks/Weidenfeld and Nicolson: 1997.

Pp. 168. \$20, £11.99

Alison Jolly

"Recreational sex and menopause were as important for our development of fire, language, art and writing as were our upright posture and large brains." Jared Diamond sets up a major claim at the start of *Why is Sex Fun?* He then discusses not only recreational sex, but the evolutionary reasons why men philander more than women, whom it pays to desert first, why men don't breast-feed even though they could (and if *in vitro* fertilization keeps producing twins, perhaps will yet do so), and many other intriguing thoughts, ending with the overgrown human penis.

The author assumes no prior knowledge. Little of the story will surprise readers of the literature, or of Diamond's wonderful *Third Chimpanzee* (HarperCollins, 1992). But it is beautifully told, with male breast-feeding introducing a primer on embryonic sex differentiation, menopause one on life history strategy, and concealed ovulation

the principle that evolution works on what is there.

Concealed ovulation and sex for fun may well have changed over time from the 'many-fathers' function, with troop males bamboozled into tolerating every infant as their own, to the opposite 'daddy-at-home' function of tribal males' paternal confidence, directed child-care and frequent sex with a loving—maybe even monogamous—wife.

Diamond's recurrent ploy is to present his case, then add, "By now you are probably objecting that..." This lets him deal with counter-arguments while flattering the reader, who actually had only a dim feeling that there must be a catch somewhere. This book is simply written, but not simply thought out—it is a good book to leave around, half-hidden, to tempt teenage offspring to become biologists.

But what of Diamond's first big claim? He makes a clear case that monogamy within a community group, sex in private, menopause and our high parental investment are unique among our close primate kin. He does not quite sum up: there should be two pages more to pull the scenario together. Does he think that our unique sexual behaviour is itself what needs explaining, or is it an explanation?

Does he see a bipedal male australopithecine striding along like one of his New Guinea trackers, head high and hands free for defence against the genocidal males of the next group? And the female finding safety and paternal care more and more from a single mate as she evolves into early *Homo*, and focuses her extended sexual attractiveness from the many to the few, in private?

If this mating system accompanied the earliest expansion of the hominid brain, it would have provided a way for extended parental care to foster the slow growth of big-brained children (a familiar thought), while allowing an almost indefinitely expanded social group, which in turn would have demanded the brainpower of an expanding neocortex.

Diamond adds a factor much less often considered than the needs of children: the value of the old. A toothless crone who still recalls the famine food she ate after the hurricane of 1910 becomes both the tribal university library and their spaceship survival manual. Are longevity and menopause not just the result of bigger brains? Perhaps they evolved simultaneously or were even prerequisites? And how does this accord with Diamond's own preference for a 'big bang' emergence of language, art and religion a mere 30,000 years ago? Perhaps he should now write his missing two-page scenario. □

Alison Jolly is in the Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544-1003, USA.

The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean

D. Karl*, R. Letelier†, L. Tupas*, J. Dore*‡, J. Christian§ & D. Hebel*

* School of Ocean and Earth Science and Technology, Department of Oceanography, University of Hawaii, Honolulu, Hawaii 96822, USA

† College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331-5503, USA

‡ Aquasearch Inc., 73-4460 Queen Kaahumanu Highway, Suite 110, Kailua-Kona, Hawaii 96740, USA

§ Department of Oceanography, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada

Seven years of time-series observations of biogeochemical processes in the subtropical North Pacific Ocean gyre have revealed dramatic changes in the microbial community structure and in the mechanisms of nutrient cycling in response to large-scale ocean-atmosphere interactions. Several independent lines of evidence show that the fixation of atmospheric nitrogen by cyanobacteria can fuel up to half of the new production. These and other observations demand a reassessment of present views of nutrient and carbon cycling in one of the Earth's largest biomes.

Biological dinitrogen (N_2) fixation—the exclusively prokaryotic metabolic process responsible for converting the most abundant but relatively inert form of N into biologically available substrates—is the dominant mechanism for introduction of N into the biosphere; it approximates the total losses from microbiological denitrification on the global scale. Consequently, data on the quantitative role of N_2 fixation in the oceans' nutrient budgets is of considerable scientific interest¹. Although the presence of N_2 -fixing microorganisms and the associated nitrogenase activity in stratified, low-nutrient oceanic habitats is well documented^{2–5}, it has been difficult to obtain reliable estimates of N_2 -fixation rates, in part owing to spatial and temporal undersampling of the marine environment^{6–10}. Most historical estimates have shown low rates of N_2 fixation relative to the total N requirement for primary production in the sea. Consequently, most oceanic biogeochemical models simply ignore the potential flux from N_2 fixation into these habitats¹¹.

New production, defined as the amount of organic matter available for export from the surface ocean at steady state¹², is thought to be constrained by the upward vertical flux of nitrate across the permanent thermocline¹³. The biologically mediated export of particulate organic matter from the surface ocean is a crucial term in the global carbon cycle because it represents a potential long-term sink for atmospheric carbon dioxide. Therefore, it is necessary to have an accurate understanding of new production mechanisms in the open-ocean regions where 70% of the total global export takes place¹⁴.

Here we summarize several independent data sets from the continuing Hawaii Ocean Time-series (HOT) research programme that collectively provide evidence for the importance of N_2 fixation as a source of new N for the pelagic ecosystem at a station that is believed to be characteristic of the subtropical North Pacific Ocean^{15,16}. Nitrogen-budget estimates suggest that N_2 fixation may supply up to half of the N required to sustain the rate of the annual particulate N (PN) export from the euphotic zone. However, our observations indicate that this relatively high percentage of N_2 -supported production may represent a transient ecosystem state reflecting either natural oceanic variability or, perhaps, an unusual state established in response to the well documented decade-long shift in North Pacific climate¹⁷.

Station ALOHA field data

Since October 1988, a comprehensive suite of hydrographical and

biogeochemical data have been collected at approximately monthly intervals to assess seasonal and interannual ocean variability at a time-series hydrostation (station ALOHA) located 100 km north of Oahu, Hawaii (22° 45' N, 158° W) in the North Pacific subtropical gyre¹⁵.

One important, unexpected source of biogeochemical variability in this region has been the aperiodic appearance of large concentrations of the phototrophic, N_2 -fixing cyanobacterium *Trichodesmium* spp.¹⁰. Long-term (>1 yr duration) increases in the abundances of *Trichodesmium* and other free-living and endosymbiotic N_2 -fixing microorganisms may be a manifestation of extended changes in the subtropical habitat¹⁸ induced by the El Niño/Southern Oscillation (ENSO). It has been proposed that this additional source of new nitrogen to the pelagic ecosystem has driven the North Pacific subtropical gyre from a nitrogen- to a phosphorus-controlled state¹⁸; we report here ecosystem data, collected over a 7-year period, that support this hypothesis.

Our evidence of substantial N_2 -supported new and export production is derived from several independent measurements and data syntheses, including: (1) *Trichodesmium* population abundances and estimates of their potential rates of biological N_2 fixation; (2) assessment of the molar N:P stoichiometries of surface-ocean dissolved and particulate matter pools and development of a one-dimensional model to calculate N and P mass balances; (3) seasonal variations in the natural ¹⁵N isotopic abundances of particulate matter exported to the deep sea and collected in bottom-moored sediment traps; and (4) observations on the secular changes in soluble reactive P (SRP), soluble non-reactive P (SNP) and dissolved organic N (DON) pools during the period of increased rates of N_2 fixation. Considered separately, none of these data sets provides conclusive evidence for the quantitative role of N_2 fixation. But when these data sets are viewed together, in the context of other time-series observations, we consider that they fully support our conclusion about the significance of N_2 fixation in the subtropical North Pacific Ocean.

Trichodesmium abundances and N_2 -fixation rates

Although N_2 fixation can occur in free-living and endosymbiotic cyanobacteria as well as in selected species of heterotrophic bacteria, the filamentous cyanobacterium *Trichodesmium* spp. is considered to be a dominant N_2 -fixing species in oligotrophic marine ecosystems^{1,9,19}. Using epifluorescence microscopy, we have quantified

the abundance of *Trichodesmium* (*T. thiebautii* and *T. erythraeum*) at station ALOHA for a period of three years and have reported a significant increase in *Trichodesmium* spp. biomass in the upper portion (0–45 m) of the water column over this observation period²⁰. Based on a conservative estimate of cell-specific N₂ fixation of 2 $\mu\text{mol N}_2$ fixed per mmol of cell N per hour for *Trichodesmium* samples previously collected in the North Pacific subtropical gyre²¹, we estimate that the *Trichodesmium* population is capable of supplying 51 (± 26) mmol N m⁻² yr⁻¹. This value is equivalent to ~50% of the PN export measured at station ALOHA for the period October 1988 to December 1995 (Table 1). Rates of N₂ fixation estimated from direct measurements of nitrogenase activity (acetylene reduction) of isolated *Trichodesmium* colonies and free trichomes collected at station ALOHA during several cruises are consistent with the rates extrapolated from the much larger data set on *Trichodesmium* abundance measurements (Table 1). These rate estimates are probably conservative because they do not include the potential contributions from other species that are known to be present at station ALOHA, most notably *Richelia intracellularis*, which is found in symbiotic associations with the diatom *Rhizosolenia* spp.⁵, and the free-living cyanobacterium *Synechococcus* spp.²².

N:P stoichiometry of dissolved and particulate material

The assessment of quantitative relationships between N and P in dissolved and particulate matter pools provides a conceptual framework for studies of organic-matter production-remineralization and nutrient limitation in the sea^{23,24}. During nutrient-sufficient conditions near maximal growth rate, the N:P molar ratio for organic-matter production in the ocean should converge on 16:1 (the Redfield ratio)²⁴. Thus attainment of the predicted Redfield N:P

stoichiometry in natural environments suggests rapid growth rate and complete nutrient saturation²⁴. Because both positive and negative deviations from this predicted cellular N:P stoichiometry occur during nutrient-limited microbial growth²⁵, it is possible to use the N:P ratio as a diagnostic ecosystem parameter.

During the first two years of the HOT programme (pre-ENSO), the mean N:P ratio for suspended particulate matter in the upper (0–100 m) water column was 15.3 (standard deviation (s.d.) = 3.1, $n = 14$), a value that was not significantly different from the Redfield prediction of 16.0 (Fig. 1). The transient, elevated N:P ratios observed during summer 1989 may have been a direct consequence of the extensive *Trichodesmium* bloom reported for this region¹⁰. Since 1991, however, we observed an increase in the molar N:P ratio of suspended particulate matter to a value greater than the expected Redfield stoichiometry (Fig. 1); that is, the mean N:P ratio for the period 1991–95 is 21.0 (s.d. = 3.5, $n = 44$). Because selected species of cyanobacteria, including *Trichodesmium*, can grow with reduced cell quotas of phosphorus¹⁰, the ecosystem drifts out of a Redfield N:P balance *en route* to phosphorus limitation as the supply of new N shifts from a limiting flux of nitrate from below the euphotic zone to the inexhaustible pool of N₂ that is dissolved in the surface waters of the ocean. This shift to N₂-supported new and export production has significant consequences for biogeochemical cycling pathways and rates. The coherent temporal pattern observed for suspended N:P, with maxima in the summer periods, is consistent with enhanced bioavailability of N relative to P resulting from increased rates of N₂ fixation during periods of maximum water-column stratification^{18,26}.

The observed N:P balance of the dissolved nutrient reservoir, and the stoichiometry of the exported particulate matter, can also be

Table 1 Selected nitrogen inventories and fluxes at station ALOHA during 1988–95

Parameter and period	Method	Value*
Inventories (mmol N m ⁻²)		
Nitrate (0–100 m) 1988–95 Winter (Jan.–Mar.) Summer (Jul.–Sept.) Annual mean	Chemiluminescence	2.6 (±4.0) 0.4 (±0.3) 1.0 (±2.2)
TDN (0–100 m) 1988–95	UV photo-oxidation/colorimetry	568 (±62)
PN (0–100 m) 1988–95	High temperature combustion/gas chromatography	32 (±9)
Fluxes (mmol N m ⁻² yr ⁻¹)		
PN production 1988–95	¹⁴ C method, assuming Redfield C:N molar ratio of 6.6	2,044 (±694)
PN export 1988–95	Free-drifting sediment traps deployed at 150 m	105 (±43)
N ₂ fixation Oct. 1988–Dec. 1992	<i>Trichodesmium</i> abundance, assuming cell-specific N ₂ -fixation rate†	51 (±26)
N ₂ fixation 1990–92	Based on HOT program acetylene reduction measurements of nitrogenase activity converted into N ₂ -fixation equivalents‡	31 (±18)
N ₂ fixation 1988–95	N:P mass balance (see Fig. 2 and text)	34 (±14)
N ₂ fixation 1992–93	¹⁵ N isotope balance (see Table 2 and text)	50
N ₂ fixation 1988–95	Range of independent estimates	31–51

TDN, total dissolved N; PN, particulate nitrogen.

*Mean values, along with our best estimates of uncertainty.

†Based on reported trichome abundances ($4.6 \times 10^4 \pm 2.3 \times 10^4$) trichomes per m³ for 0–45 m depth stratum) and elemental composition of *Trichodesmium* (9.6 ng N per trichome) for station ALOHA¹⁰, and assuming a conservative, average specific rate of N₂ fixation of 2 $\mu\text{mol N}_2$ fixed per mmol cell N per hr (ref. 21).

‡Based on acetylene reduction rates of *Trichodesmium* spp. measured for colonies collected at station ALOHA in June 1990 and February 1991, and in colonies and single trichomes collected at station ALOHA in August 1992 converted to N₂-fixation rates assuming an ethylene-to-N₂ ratio of 3 (ref. 40) and extrapolated total N₂ fixation using temporal distribution of *Trichodesmium* biomass measured at station ALOHA²⁰

explained by the growth of N_2 -fixing microorganisms in the pelagic environment of the North Pacific. The average total dissolved N: total dissolved P pool ratio (TDN:TDP) for our 7-year data set is much greater than the expected Redfield ratio, indicating an accumulation of N relative to P (Fig. 1). The average dissolved N:P ratios for the depth intervals 0–100 m and 100–200 m show coherent temporal trends that are distinct from those observed below the euphotic zone (Fig. 1). The large difference in the TDN:TDP ratios in the upper water column (0–200 m) at station ALOHA, compared to deeper values (200–500 m), indicates that there must be a net contemporaneous production of dissolved fixed N in the surface waters, relative to the large nutrient reservoir below the euphotic zone. During episodes of enhanced vertical mixing, the surface-water TDN:TDP ratio returns to a near-Redfield balance as

deeper waters are injected into the euphotic zone (for example, in spring 1990, 1992 and 1995; Fig. 1). Conversely, prolonged stratification, as might be expected during frequent or intense ENSO periods¹⁸, affects the N:P stoichiometry of both dissolved and particulate matter pools in the surface (0–100 m) waters, and forces the ecosystem out of the expected Redfield balance (for example, March 1992–October 1994; Fig. 1).

The N:P stoichiometry of exported particulate matter collected at the 150-m reference level is consistent with the trends already discussed for suspended and dissolved matter pools. Low N:P values that were observed following a deep mixing event during HOT-14 (spring 1990), and the sustained higher than Redfield-ratio export, recorded especially during the extended ENSO period, are among the coherent trends (Fig. 1).

Figure 1 Time series of N and P analyses of dissolved and particulate matter, expressed as N:P (mol:mol) ratios, for dissolved matter (upper panel), suspended particulate matter (centre panel) and exported particulate matter (lower panel). The upper panel shows the 3-point running mean N:P ratios for 0–100 m (●), 100–200 m (■) and 200–500 m (▲) portions of the water column. These ratios were determined by integration (trapezoid rule) of discrete total dissolved nitrogen (TDN) and total dissolved phosphorus (TDP) measurements ($n = 4-8$) for each depth region, and calculation of a single TDN:TDP ratio from these inventory estimations. The centre panel shows the 3-point running mean (± 1 standard deviation, s.d.) for the average suspended particulate N:P ratio measured in the upper portion (0–100 m) of the water column on each cruise (depth-integrated particulate N $\div$ depth-integrated particulate P). The lower panel shows the 3-point running mean (± 1 s.d.) for the average N:P ratio of the sediment-trap-collected particulate matter at the 150-m reference depth. Water samples were collected in polyvinylchloride bottles using standard shipboard procedures¹⁵. Subsamples for TDN and TDP concentrations were treated with high-intensity ultraviolet light, followed by automated colorimetric analysis³⁹. Subsamples for suspended particulate N analyses were screened through 202- μ m Nitex then filtered through combusted, Whatman (GF/F) glass fibre filters followed by high-temperature combustion and detection of CO_2 and N_2 (ref. 35). Suspended particulate P was measured by high-temperature sample ashing and autoanalysis⁴⁰. Free-drifting sediment traps³⁵ deployed at 150 m for an ~ 3 -day period during each cruise were used to collect sinking particulate matter for the determination of export fluxes and for N and P elemental analyses, as above, except that 335- μ m Nitex pre-screening was used. The Redfield ratio (N : P = 16) is represented by a dashed line in all three panels.

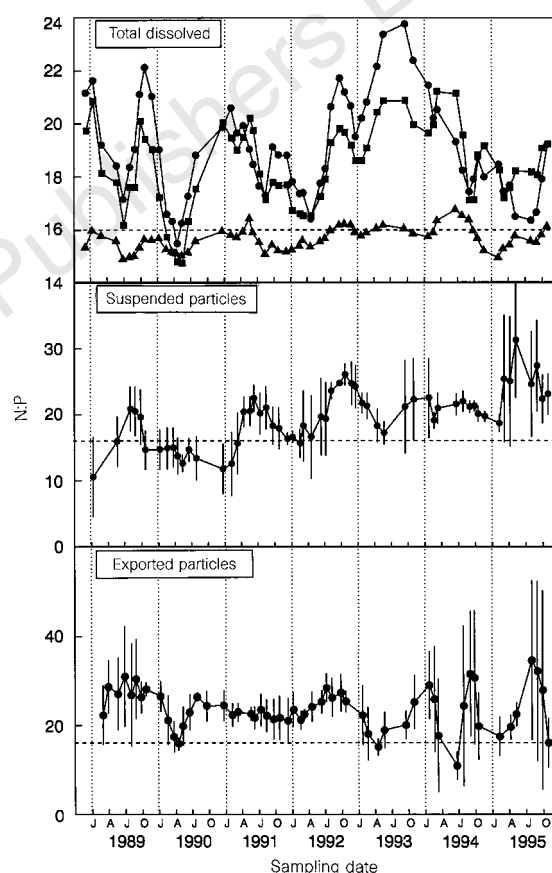
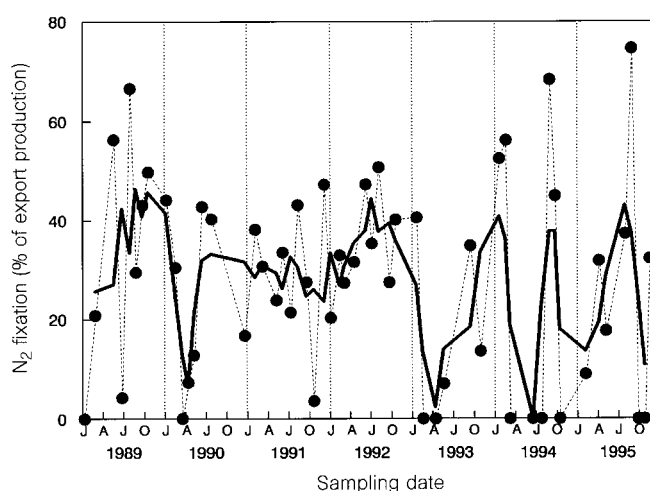


Figure 2 Role of N_2 fixation as a source of new nitrogen as determined by the N–P mass-balance model (see text). The data shown are the monthly estimates of N_2 fixation, expressed as a percentage of particulate nitrogen (PN) export for the 7-year HOT programme data set. The solid line is the 3-point running mean value. Cruises where the calculated percentage was < 0 (that is, where $(N : P)_{out} < (N : P)_{up}$) were treated as 0% N_2 fixation for the purpose of calculating the 3-point running mean. The 7-year average N_2 fixation at station ALOHA based on the N–P mass balance model is 32% of PN flux.



Below the euphotic zone, the exported organic matter is solubilized and remineralized by a variety of processes. It is important to reiterate that the relatively large intermediate-depth (200–500 m) reservoir of dissolved nutrients has an average N:P ratio that is distinct from either the surface ocean pools or the sinking particulate matter collected at station ALOHA. However, the N_2 -fixation signatures that we now observe in the HOT programme data sets may be the result of a transient N–P cycle decoupling, rather than being representative of the long-term (>100-year) steady state. Otherwise, we might expect the N:P ratios in the intermediate and shallow pools to converge because surface ocean export largely fuels nutrient regeneration at depth.

A simple two-component N source model was constructed to assess the potential role of N_2 fixation using an upper water column N–P mass-balance consideration. The operative assumption is that the only new P source term, as SRP, is due to upward eddy diffusion, and that this process and N_2 fixation are the two potential source terms for new N. When considering the total pool of organic matter (dissolved plus particulate), the N:P ratio in the upper water column (0–100 m) at station ALOHA is relatively constant (mean total N : total P = 18.9, s.d. = 2.4, $n = 59$) during our observation period (October 1988 to December 1995). Consequently an assessment of the role of N_2 fixation can be made by considering the N:P ratios of various potential sources and sinks as:

$$N_{\text{fix}}/[N_{\text{fix}} + N_{\text{up}}] = [(N:P)_{\text{out}} - (N:P)_{\text{up}}]/(N:P)_{\text{out}}$$

where N_{fix} and N_{up} are the rates of N_2 fixation and upward eddy

diffusion of N, respectively, and $(N:P)_{\text{out}}$ and $(N:P)_{\text{up}}$ are the N:P ratios measured for particles exported from the euphotic zone and for the upward eddy-diffusive nutrient flux, respectively. Based on the results from station ALOHA (Fig. 1), that is, the mean $(N:P)_{\text{out}} = 23.3$ (s.d. = 9.2, $n = 56$) and the mean $(N:P)_{\text{up}} = 15.6$ (s.d. = 0.6, $n = 61$), we estimate that 32% of the particulate N export, or $34 \text{ mmol N m}^{-2} \text{ yr}^{-1}$, is supported by N_2 fixation (Fig. 2 and Table 1). Significant temporal variations were observed in the estimated percentage of PN export from N_2 fixation, with values ranging from 0% following deep winter mixing events (for example, 1990, 1993, 1994) to >50% during periods of extended water column stratification (for example, 1989, 1992, 1994, 1995; Fig. 2). We also observed large interannual variations in the implied role of N_2 fixation, with a sustained 2-year maximum during the extended ENSO period (1991–92) to a distinct minimum in 1994 (Fig. 2). These seasonal and interannual variations in the rates of N_2 fixation are inextricably coupled to upper-ocean physical conditions and, ultimately, to ocean–atmosphere interactions.

We also observed a systematic change in the inventory of DON in the upper water column (0–100 m) that was especially evident during the 1991–92 ENSO event. The nitrate inventories were relatively constant and were, on average, <1% of the TDN pool. There was no evidence of a winter-time nitrate injection as we have observed in non-ENSO years (Table 1). Nevertheless, there was a dramatic increase in DON in surface waters beginning in the summer of 1991 which was sustained until the end of 1992 (Fig. 3). A similar,

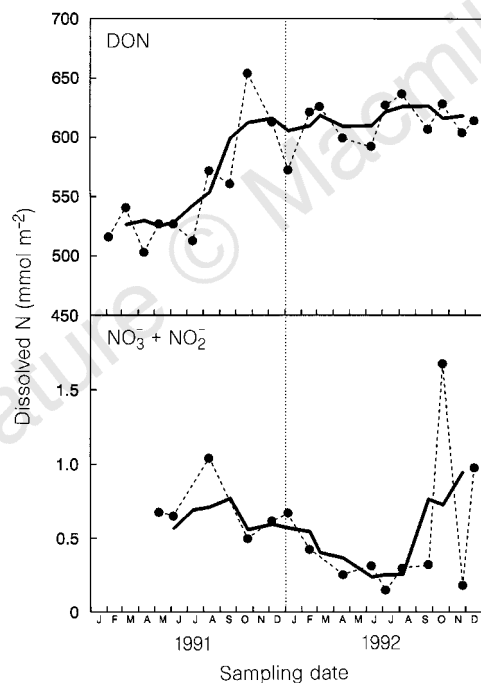


Figure 3 Dissolved nitrogen pool dynamics for the upper portion of the water column (0–100 m) at station ALOHA during the 1991–92 ENSO event. Top, dissolved organic nitrogen (DON) pool inventories; bottom, nitrate plus nitrite pool inventories. The solid line is the 3-point running mean value for each determination. Note the concentration scale differences; the majority of the total dissolved nitrogen (TDN) pool (>99%) is DON. $[\text{NO}_3^- + \text{NO}_2^-]$ was measured by chemiluminescence⁴¹, except for July 1991 and March 1992 when it was measured using standard autoanalyser techniques. DON was measured by difference ($\text{TDN} - [\text{NO}_3^- + \text{NO}_2^-]$) following high-intensity ultraviolet light treatment³⁹. Each data point represents the 0–100 m depth-integrated inventory of $[\text{NO}_3^- + \text{NO}_2^-]$ or DON measured on the approximately monthly cruises, generally from 6–7 individual water depths.

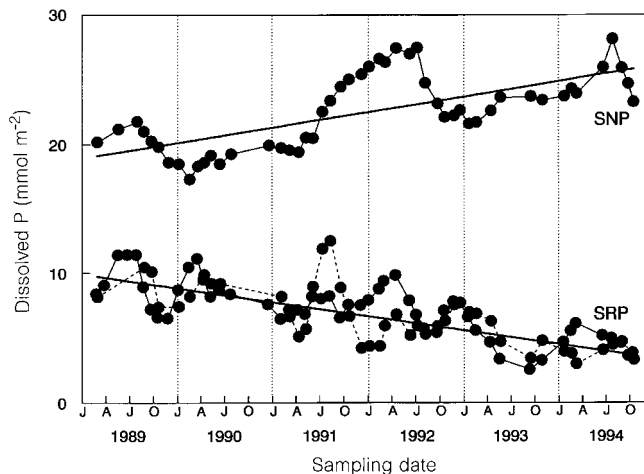


Figure 4 Dissolved phosphorus pool dynamics at station ALOHA for the period 1989–94, including the period corresponding to the extended 1991–92 ENSO event. The concentration of soluble reactive phosphorus (SRP) was measured both by a standard autoanalyser technique (data points connected by solid line) and by the high-sensitivity magnesium-induced co-precipitation (MAGIC)⁴² technique (data points connected by dashed line). Soluble non-reactive phosphorus (SNP) was measured by difference ($\text{TDP} - \text{SRP}$) following high-intensity ultraviolet light treatment³⁹. Each data point represents the 0–100 m depth-integrated inventory of SRP or SNP measured on the cruises, generally from 6–7 individual water depths. The model I linear regression fit (bold line) for the MAGIC-SRP data set is: $\text{SRP} (\text{mmol m}^{-2}) = 9.95 - (0.003 \times \text{days since 1 Jan. 1989})$. The model I linear regression fit (bold line) for the SNP data set is: $\text{SNP} (\text{mmol m}^{-2}) = 18.98 + (0.003 \times \text{days since 1 Jan. 1989})$.

Table 2 Nitrogen-isotope determinations of selected particulate pools

Sample source	$\delta^{15}\text{N}^*$ (‰)	Estimated contribution from N_2 (% of total N) [†]
<i>Trichodesmium</i> spp.		
Laboratory culture, fixed N-free medium	-0.38	100
Field collections (HOT-9, 0–1 m)	0.41	93
Bottom-moored sediment trap [‡]		
Winter export pulse (Feb.)	4.83	20
Summer export pulse (Jul.)	1.53	75
Annual mean	3.10	48

All data measured in the Hawaii Ocean Time-series programme.

*Relative to air N_2 .

[†]Based on a two-source model: $\delta^{15}\text{N} = (\delta^{15}\text{N}_a)f + (\delta^{15}\text{N}_b)(1-f)$, where f is the decimal fractional equivalent, N_a and N_b are theoretical values for N_2 ($\delta^{15}\text{N} = 0$) and NO_3^- ($\delta^{15}\text{N} = 6.5$) based on values reported for the North Pacific Ocean (ref. 32).

[‡]A continuously recording, bottom-moored time-series sediment trap^{35,44} with traps positioned at 800, 1,500, 2,800 and 4,000 m below the surface was used to collect the sinking particulate materials for the measurement of N isotope ratios. Particulate matter from the 1,500-m trap was concentrated onto Whatman GF/F filters followed by sealed-tube combustion with Cu and CuO to N_2 , and gas analysis using a Finnigan MAT-252 isotope-ratio mass spectrometer.

but transient, increase in DON was reported for the August 1989 *Trichodesmium* bloom at station ALOHA¹⁰. *Trichodesmium* is known to release up to 50% of the N_2 fixed as dissolved nitrogen²⁷, in particular as the amino acid glutamate²⁸. Compared to the 1990 annual mean DON of $474 \text{ mmol N m}^{-2}$ (s.d. = 65; $n = 8$), the 1991 and 1992 values (mean = 552, s.d. = 49, $n = 10$ and mean = 612, s.d. = 19, $n = 11 \text{ mmol N m}^{-2}$, respectively) represent a sustained accumulation of $69 \text{ mmol N m}^{-2} \text{ yr}^{-1}$ during this ENSO event. With an average C:N ratio of ~ 20 for the dissolved organic matter at station ALOHA²⁹, this DON accumulation could represent a net carbon sequestration of $\sim 1.4 \text{ mol C m}^{-2} \text{ yr}^{-1}$ during the 1991–92 ENSO. This value exceeds the estimated $0.8\text{--}1.0 \text{ mol C m}^{-2} \text{ yr}^{-1}$ annual air-to-sea flux of carbon dioxide for this region³⁰, and emphasizes the potential role of N_2 fixation in the C as well as the N cycle at station ALOHA. These DON accumulation results are consistent with the conceptual model previously presented for the ENSO-induced, N_2 -fixation-mediated alternation from N- to P-limitation at station ALOHA¹⁸.

Nitrogen-isotope composition of sinking particles

The measurement of the natural abundance of ^{15}N in particulate matter can be used as a diagnostic tracer of N sources for organic production. The nitrogen isotope ratio ($^{15}\text{N}/^{14}\text{N}$), expressed in the conventional $\delta^{15}\text{N}$ (‰) notation³¹, is potentially quite useful in our study because there is a large difference in N isotope ratio between the two different 'new' N sources ($\text{N}_2 = 0\text{‰}$ versus $\text{NO}_3^- = 6\text{--}7\text{‰}$)³². Although the exact $\delta^{15}\text{N}$ value of particulate matter is a complex result of the $\delta^{15}\text{N}$ value of the N source and the degree of isotopic fractionation during assimilation, the extremely low $\delta^{15}\text{N}$ values reported for *Trichodesmium* (approximately equal to the $\delta^{15}\text{N}$ of atmospheric N_2)³³ provide an unequivocal assessment of N_2 fixation in our study area. Time-series measurements of the $\delta^{15}\text{N}$ of sinking particulate matter collected using a bottom-moored sediment trap deployed near station ALOHA reveal a systematic seasonal flux pattern that is consistent with our model of nitrate-based export production following periods of winter mixing and N_2 -based export production during stratified summer conditions (Table 2). From the results of a simple two-source model, we estimate that N_2 fixation supplies 48% of the total annual PN flux collected at the 1,500 m reference depth (Table 2). This value is

equivalent to $50 \text{ mmol N m}^{-2} \text{ yr}^{-1}$ of N_2 -based new production, if a similar proportion holds for particulate matter exported at the base of the euphotic zone.

Dissolved-phosphorus dynamics

The dynamics of the dissolved phosphorus pool in the upper water column are also consistent with the suggested role of N_2 fixation at station ALOHA. For example, the most readily available pool of phosphorus, SRP, decreased by more than 50% between 1989 and 1994 at an average rate of $1.1 \text{ mmol P m}^{-2} \text{ yr}^{-1}$, presumably in response to the shift from N to P limitation caused by increased N_2 fixation (Fig. 4). By comparison, the near-surface (0–100 m) SNP pool, consisting largely of dissolved organic P compounds, increased over this same period at an average rate that was indistinguishable from the SRP decrease (that is, $1.1 \text{ mmol P m}^{-2} \text{ yr}^{-1}$). We interpret this shift from SRP to SNP to be the result of the production and accumulation of refractory P compounds that are not readily available to the plankton community. Direct measurements of microbial uptake of selected SNP compounds in Hawaiian waters indicate that SNP assimilation rates are an order of magnitude lower than those measured for orthophosphate (that is, SRP) when added at equivalent concentrations³⁴. Extended periods of N_2 fixation are expected to favour the selective removal of SRP and the selective accumulation of SNP, exactly the pattern that we observe at station ALOHA (Fig. 4).

During the period when SNP was accumulating in the upper water column (0–100 m), there was a corresponding increase in DON that was especially evident during the ENSO period (Fig. 3). The molar N:P ratio of the accumulating dissolved pool during 1991–92 is 63, nearly four times greater than the Redfield ratio. We interpret these coupled N–P nutrient dynamics (SRP decrease, SNP and DON increases and non-Redfield dissolved-matter accumulation) at station ALOHA to be the result of an uncoupling in the input of N versus P into the pelagic ecosystem. This uncoupling is consistent with an increase in cyanobacterial N_2 fixation over the period of study. It is also informative to point out that during the period of DON accumulation, total autotrophic production (measured using ^{14}C techniques) increased without a concomitant increase in particulate matter¹⁸ while the sediment trap PN export decreased relative to the 7-year average^{18,35}. These results suggest that N_2 -supported new production pathways may also alter the mechanics of nutrient export from a predominantly particulate-matter export to one where dissolved-matter export processes may become more important.

Station ALOHA nitrogen budget

We have compiled a budget for our time-series station based on various measured and derived fluxes of nitrogen. From these independent analyses, we conclude that N_2 fixation may contribute up to half of the N required to sustain total annual export in the oligotrophic North Pacific Ocean (Table 1). There are at least two significant consequences of these results. First, it seems that the process of N_2 fixation is an important source of 'new' N (as defined by Dugdale and Goering³⁶) at station ALOHA and, probably, for the oligotrophic North Pacific subtropical gyre as a whole, especially during the 1991–92 ENSO period. Whether this temporal coherence with ENSO is a direct cause-and-effect relationship¹⁸ cannot be determined from the available data sets. Because nitrogenase requires Fe as a cofactor³⁷, the availability of this and other essential trace nutrients may also exert a control on ambient rates of N_2 fixation. Direct measurements of dissolved Fe at station ALOHA indicate the presence of non-limiting concentrations (0.2–1.0 nM; ref. 38 and C. Measures, personal communication). The potential interactions of Fe deposition at the surface with N_2 -fixation rates and the attendant biogeochemical processes discussed here provides a basis on which to examine climate regimes. Second, it seems that phosphorus pool dynamics may ultimately control new and export

production at station ALOHA, and probably in the subtropical North Pacific Ocean as a whole, during periods of enhanced N_2 fixation. This revised view of the control of organic-matter production processes in possibly the largest continuous biome on Earth may require other changes in our assessment of global biogeochemical cycles. For example, if our conclusion about the role of cyanobacterial N_2 fixation in new and export production is correct, we predict that new N inputs to these low-nutrient habitats may be more dependent upon the relaxation of upper-ocean mixing (selection for N_2 -fixing microorganisms) than on increased atmosphere-forced mixing as is presently considered in new-production models. Furthermore, our results suggest a previously undetected seasonal variability in sources of new N, as we have documented with the $\delta^{15}N$ data set derived from sediment-trap data. This altered view of biogeochemical dynamics in the subtropical North Pacific Ocean may have a profound influence on how we model ecosystem processes, including the potential effects of natural or human-induced environmental change. $\square$

Received 4 June; accepted 27 June 1997.

- Capone, D. G., Zehr, J. P., Paerl, H. W., Bergman, B. & Carpenter, E. J. *Trichodesmium*, a globally significant marine cyanobacterium. *Science* **276**, 1221–1229 (1997).
- Dugdale, R. C., Menzel, D. W. & Ryther, J. H. Nitrogen fixation in the Sargasso Sea. *Deep-Sea Res.* **7**, 297–300 (1961).
- Goering, J. J., Dugdale, R. C. & Menzel, D. W. Estimates of rates of nitrogen uptake by *Trichodesmium* sp. in the tropical Atlantic Ocean. *Limnol. Oceanogr.* **11**, 614–620 (1966).
- Carpenter, E. J. Nitrogen fixation by *Oscillatoria* (*Trichodesmium*) *thiebautii* in the southwestern Sargasso Sea. *Deep-Sea Res.* **20**, 285–288 (1973).
- Mague, T. H., Weare, N. M. & Holm-Hansen, O. Nitrogen fixation in the North Pacific Ocean. *Mar. Biol.* **24**, 109–119 (1974).
- Carpenter, E. J. & Price, C. C. Nitrogen fixation, distribution, and production of *Oscillatoria* (*Trichodesmium*) spp. in the western Sargasso and Caribbean Seas. *Limnol. Oceanogr.* **22**, 60–72 (1977).
- Howarth, R. W., Marino, R., Lane, J. & Cole, J. J. Nitrogen fixation in freshwater, estuarine, and marine ecosystems. 1. Rates and importance. *Limnol. Oceanogr.* **33**, 669–687 (1988).
- Capone, D. G. & Carpenter, E. J. Nitrogen fixation in the marine environment. *Science* **217**, 1140–1142 (1982).
- Carpenter, E. J. & Romans, K. Major role of the cyanobacterium *Trichodesmium* in nutrient cycling in the North Atlantic Ocean. *Science* **254**, 1356–1358 (1991).
- Karl, D. M., Letelier, R., Hebel, D. V., Bird, D. F. & Winn, C. D. in *Marine Pelagic Cyanobacteria: Trichodesmium and Other Diazotrophs* (eds Carpenter, E. J., Capone, D. G. & Rueter, J. G.) 219–237 (Kluwer Academic, Dordrecht, 1992).
- Altabet, M. A., Robinson, A. R. & Walstad, L. J. A model for the vertical flux of nitrogen in the upper ocean: simulating the alteration of isotopic ratios. *J. Mar. Res.* **44**, 203–225 (1986).
- Eppley, R. W. & Peterson, B. J. Particulate organic flux and planktonic new production in the deep ocean. *Nature* **282**, 677–680 (1979).
- Lewis, M. R., Harrison, W. G., Oakey, N. S., Hebert, D. & Platt, T. Vertical nitrate fluxes in the oligotrophic ocean. *Science* **234**, 870–873 (1986).
- Martin, J. H., Knauer, G. A., Karl, D. M. & Broenkow, W. W. VERTEX: carbon cycling in the northeast Pacific. *Deep-Sea Res.* **34**, 267–286 (1987).
- Karl, D. M. & Lukas, R. The Hawaii Ocean Time-series (HOT) Program: Background, rationale and field implementation. *Deep-Sea Res.* **43**, 129–156 (1996).
- Karl, D. M. HOT and the North Pacific gyre. *Nature* **378**, 21–22 (1995).
- Trenberth, K. E. & Hurrell, J. W. Decadal atmosphere-ocean variations in the Pacific. *Clim. Dyn.* **9**, 303–319 (1994).
- Karl, D. M. *et al.* Ecosystem changes in the North Pacific subtropical gyre attributed to the 1991–92 El Niño. *Nature* **373**, 230–234 (1995).
- Carpenter, E. J., Capone, D. G. & Rueter, J. G. (eds) *Marine Pelagic Cyanobacteria: Trichodesmium and other Diazotrophs* (Kluwer Academic, Dordrecht, 1992).
- Letelier, R. M. & Karl, D. M. Role of *Trichodesmium* spp. in the productivity of the subtropical North Pacific Ocean. *Mar. Ecol. Prog. Ser.* **133**, 263–273 (1996).
- Carpenter, E. J. Physiology and ecology of marine planktonic *Oscillatoria* (*Trichodesmium*). *Mar. Biol. Lett.* **4**, 69–85 (1983).
- Mitsui, A. *et al.* Strategy by which nitrogen-fixing unicellular cyanobacteria grow photoautotrophically. *Nature* **323**, 720–722 (1986).
- Redfield, A. C., Ketchum, B. H. & Richards, F. A. in *The Sea: Ideas and Observations on Progress in the Study of the Seas Vol. 2* (ed. Hill, M. N.) 26–77 (Wiley Interscience, New York, 1963).
- Goldman, J. C., McCarthy, J. J. & Peavey, D. G. Growth rate influence on the chemical composition of phytoplankton in oceanic waters. *Nature* **279**, 210–215 (1979).
- Sakshaug, E. & Holm-Hansen, O. Chemical composition of *Skeletonema costatum* (Grev.) Cleve and *Pavlova* (*Monochrysis*) *lutheri* (Droop) Green as a function of nitrate-, phosphate-, and iron-limited growth. *J. Exp. Mar. Biol. Ecol.* **29**, 1–34 (1977).
- Carpenter, E. J. & Price, C. C. Marine *Oscillatoria* (*Trichodesmium*): Explanation for aerobic nitrogen fixation without heterocysts. *Science* **191**, 1278–1280 (1976).
- Glibert, P. M. & Bronk, D. A. Release of dissolved organic nitrogen by marine diazotrophic cyanobacteria. *Trichodesmium* spp. *Appl. Environ. Microbiol.* **60**, 3996–4000 (1994).
- Capone, D. C., Ferrier, M. D. & Carpenter, E. J. Amino acid cycling in colonies of the planktonic marine cyanobacterium *Trichodesmium thiebautii*. *Appl. Environ. Microbiol.* **60**, 3989–3995 (1994).
- Karl, D. M., Tien, G., Dore, J. & Winn, C. D. Total dissolved nitrogen and phosphorus concentrations at U.S.-JGOFS Station ALOHA: Redfield reconciliation. *Mar. Chem.* **41**, 203–208 (1993).
- Winn, C. D., Mackenzie, F. T., Carrillo, C. J., Sabine, C. L. & Karl, D. M. Air-sea carbon dioxide exchange in the North Pacific subtropical gyre: Implications for the global carbon budget. *Glob. Biogeochem. Cycles* **8**, 157–163 (1994).
- Owens, N. J. P. Natural variations in ^{15}N in the marine environment. *Adv. Mar. Biol.* **24**, 389–451 (1987).
- Wada, E. & Hattori, A. *Nitrogen in the Sea: Forms, Abundances, and Rate Processes* (CRC, Boca Raton, Florida, 1991).
- Wada, E. in *Isotope Marine Chemistry* (eds Goldberg, E. D., Horibe, Y. & Saruhashi, K.) 375–398 (Uchida Rokakuho, Tokyo, 1980).
- Björkman, K. & Karl, D. M. Bioavailability of inorganic and organic phosphorus compounds to natural assemblages of microorganisms in Hawaiian coastal waters. *Mar. Ecol. Prog. Ser.* **111**, 265–273 (1994).
- Karl, D. M. *et al.* Seasonal and interannual variability in primary production and particle flux at Station ALOHA. *Deep-Sea Res.* **43**, 539–568 (1996).
- Dugdale, R. C. & Goering, J. J. Uptake of new and regenerated forms of nitrogen in primary productivity. *Limnol. Oceanogr.* **12**, 196–206 (1967).
- Rueter, J. G., Hutchins, D. A., Smith, R. W. & Unsworth, N. L. in *Marine Pelagic Cyanobacteria: Trichodesmium and other Diazotrophs* (eds Carpenter, E. J., Capone, D. G. & Rueter, J. G.) 289–306 (Kluwer Academic, Dordrecht, 1992).
- Tersol, V., Vink, S., Yuan, J. & Measures, C. I. Variations in iron, aluminum and beryllium concentrations in surface waters at Station ALOHA. *Eos* **76**, OS65 (1996).
- Armstrong, F. A., Williams, P. M. & Strickland, J. D. H. Photo-oxidation of organic matter in seawater by ultraviolet radiation, analytical and other applications. *Nature* **211**, 481–483 (1966).
- Karl, D. M., Dore, J. E., Hebel, D. V. & Winn, C. in *Marine Particles: Analysis and Characterization* (eds Hurd, D. C. & Spencer, D. W.) 71–77 (Am. Geophys. Un., Washington DC, 1991).
- Garside, C. A chemiluminescent technique for the determination of nanomolar concentrations of nitrate and nitrite in seawater. *Mar. Chem.* **11**, 159–167 (1982).
- Karl, D. M. & Tien, G. MAGIC: A sensitive and precise method for measuring dissolved phosphorus in aquatic environments. *Limnol. Oceanogr.* **37**, 105–116 (1992).
- Stewart, W. D. P., Fitzgerald, G. P. & Burris, R. H. *In situ* studies on N_2 fixation using the acetylene reduction technique. *Proc. Natl Acad. Sci. USA* **58**, 2071–2078 (1967).
- Honjo, S. & Doherty, K. W. Large aperture time-series sediment traps: design objectives, construction and application. *Deep-Sea Res.* **35**, 133–149 (1988).

Acknowledgements. We thank the HOT Program scientists and staff, and the officers and crew members of the research vessels used in our field work for making this study possible, and E. Laws for comments on the manuscript. This work was supported in part by the National Science Foundation.

Correspondence should be addressed to D.K. (email:dkarl@soest.hawaii.edu). All HOT programme data sets, including the detailed sampling and analytical protocols, are available at <http://hahana.soest.hawaii.edu>.

The complete genome sequence of the gastric pathogen *Helicobacter pylori*

Jean-F. Tomb*, Owen White*, Anthony R. Kerlavage*, Rebecca A. Clayton*, Granger G. Sutton*, Robert D. Fleischmann*, Karen A. Ketchum*, Hans Peter Klenk*, Steven Gill*, Brian A. Dougherty*, Karen Nelson*, John Quackenbush*, Lixin Zhou*, Ewen F. Kirkness*, Scott Peterson*, Brendan Loftus*, Delwood Richardson*, Robert Dodson*, Hanif G. Khalak*, Anna Glodek*, Keith McKenney*, Lisa M. Fitzgerald*, Norman Lee*, Mark D. Adams*, Erin K. Hickey*, Douglas E. Berg†, Jeanine D. Gocayne*, Teresa R. Utterback*, Jeremy D. Peterson*, Jenny M. Kelley*, Matthew D. Cotton*, Janice M. Weidman*, Claire Fujii*, Cheryl Bowman*, Larry Watthey*, Erik W. Ilin‡, William S. Hayes§, Mark Borodovsky§, Peter D. Karp||, Hamilton O. Smith¶, Claire M. Fraser & J. Craig Venter*

* The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850, USA

† Department of Molecular Biology, School of Medicine, Washington University St Louis, 660 S. Euclid Avenue, St Louis, Missouri 63110, USA

‡ Department of Biochemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden

§ School of Biology, Georgia Tech, Atlanta, Georgia 30332, USA

|| SRI International, Artificial Intelligence Center, 333 Ravenswood Avenue, Menlo Park, California 94025, USA

¶ Department of Molecular Biology and Genetics, School of Medicine, Johns Hopkins University, 725 N. Wolfe Street, Baltimore, Maryland 21205, USA

***Helicobacter pylori*, strain 26695, has a circular genome of 1,667,867 base pairs and 1,590 predicted coding sequences. Sequence analysis indicates that *H. pylori* has well-developed systems for motility, for scavenging iron, and for DNA restriction and modification. Many putative adhesins, lipoproteins and other outer membrane proteins were identified, underscoring the potential complexity of host-pathogen interaction. Based on the large number of sequence-related genes encoding outer membrane proteins and the presence of homopolymeric tracts and dinucleotide repeats in coding sequences, *H. pylori*, like several other mucosal pathogens, probably uses recombination and slipped-strand mispairing within repeats as mechanisms for antigenic variation and adaptive evolution. Consistent with its restricted niche, *H. pylori* has a few regulatory networks, and a limited metabolic repertoire and biosynthetic capacity. Its survival in acid conditions depends, in part, on its ability to establish a positive inside-membrane potential in low pH.**

For most of this century the cause of peptic ulcer disease was thought to be stress-related and the disease to be prevalent in hyperacid producers. The discovery¹ that *Helicobacter pylori* was associated with gastric inflammation and peptic ulcer disease was initially met with scepticism. However, this discovery and subsequent studies on *H. pylori* have revolutionized our view of the gastric environment, the diseases associated with it, and the appropriate treatment regimens².

Helicobacter pylori is a micro-aerophilic, Gram-negative, slow-growing, spiral-shaped and flagellated organism. Its most characteristic enzyme is a potent multisubunit urease³ that is crucial for its survival at acidic pH and for its successful colonization of the gastric environment, a site that few other microbes can colonize². *H. pylori* is probably the most common chronic bacterial infection of humans, present in almost half of the world population². The presence of the bacterium in the gastric mucosa is associated with chronic active gastritis and is implicated in more severe gastric diseases, including chronic atrophic gastritis (a precursor of gastric carcinomas), peptic ulceration and mucosa-associated lymphoid tissue lymphomas². Disease outcome depends on many factors, including bacterial genotype, and host physiology, genotype and dietary habits^{4,5}. *H. pylori* infection has also been associated with persistent diarrhoea and increased susceptibility to other infectious diseases⁶.

Because of its importance as a human pathogen, our interest in its biology and evolution, and the value of complete genome sequence information for drug discovery and vaccine development, we have

Table 1 Genome features

General	
Coding regions (91.0%)	
Stable RNA (0.7%)	
Non-coding repeats (2.3%)	
Intergenic sequence (6.0%)	
RNA	
Ribosomal RNA	Coordinates
23S-5S	445,306-448,642 bp
23S-5S	1,473,557-1,473,919 bp
16S	1,209,082-1,207,584 bp
16S	1,511,138-1,512,635 bp
5S	448,041-448,618 bp
Transfer RNA	
36 species (7 clusters, 12 single genes)	
Structural RNA	
1 species (ssrD)	629,845-630,124 bp
DNA	
Insertion sequences	
IS605 13 copies (5 full-length, 8 partial)	
IS606 4 copies (2 full-length, 2 partial)	
Distinct G + C regions	Associated genes
region 1 (33% G + C) 452-479 kb	IS605, 5SRNA and repeat 7; <i>virB4</i>
region 2 (35% G + C) 539-579 kb	cag PAI (Fig. 4)
region 3 (33% G + C) 1,049-1,071 kb	IS605, 5SRNA and repeat 7
region 4 (43% G + C) 1,264-1,276 kb	β and β' RNA polymerase, EF-G (fusA)
region 5 (33% G + C) 1,590-1,602 kb	two restriction/modification systems
Coding sequences	
1,590 coding sequences (average 945 bp)	
1,091 identified database match	
499 no database match	

sequenced the genome of a representative *H. pylori* strain by the whole-genome random sequencing method as described for *Haemophilus influenzae*⁷, *Mycoplasma genitalium*⁸ and *Methanococcus jannaschii*⁹.

General features of the genome

Genome analysis. The genome of *H. pylori* strain 26695 consists of a circular chromosome with a size of 1,667,867 base pairs (bp) and average G + C content of 39% (Figs 1 and 2). Five regions within the genome have a significantly different G + C composition (Table 1 and Fig. 1). Two of them contain one or more copies of the insertion sequence IS605 (see below) and are flanked by a 5S ribosomal RNA sequence at one end and a 521 bp repeat (repeat 7) near the other. These two regions are also notable because they contain genes involved in DNA processing and one contains 2 orthologues of the *virB4/ptl* gene, the product of which is required for the transfer of oncogenic T-DNA of *Agrobacterium* and the secretion of the pertussis toxin by *Bordetella pertussis*¹⁰. Another region is the *cag* pathogenicity island (PAI), which is flanked by 31-bp direct repeats, and appears to be the product of lateral transfer¹¹.

RNA and repeat elements. Thirty-six tRNA species were identified using tRNAscan-SE¹². These are organized into 7 clusters plus 12 single genes. Two separate sets of 23S–5S and 16S ribosomal RNA (rRNA) genes were identified, along with one orphan 5S gene and one structural RNA gene (Table 1). Associated with each of the two 23S–5S gene clusters is a 6-kilobase (kb) repeat containing a possible operon of 5 ORFs that have no database matches.

Eight repeat families (>97% identity) varying in length from 0.47 to 3.8 kb were found in the chromosome (Figs 1 and 2). Members of repeat 7 are found in intergenic regions, while the others are associated with coding sequences and may represent gene duplications. Repeats 1, 2, 3 and 6 are associated with genes that encode outer-membrane proteins (OMP) (Fig. 3).

Two distinct insertion sequence (IS) elements are present. There are five full-length copies of the previously described IS605^{11,13} and two of a newly discovered element designated IS606. In addition, there are eight partial copies of IS605 and two partial copies of IS606. Both elements encode two divergently transcribed transposases (TnpA and TnpB). IS606 has less than 50% nucleotide identity with IS605 and the IS606 transposases have 29% amino-acid identity with their IS605 counterpart. Both copies of the IS606 TnpB may be non-functional owing to frameshifts.

Origin of replication. As a typical eubacterial origin of replication was not identified¹⁴, we arbitrarily designated basepair one at the start of a 7-mer repeat, (AGTGATT)₂₆, that produces translational stops in all reading frames, as this repeated DNA is unlikely to contain any coding sequence.

Open reading frames. One thousand five hundred and ninety predicted coding sequences were identified. They were searched against a non-redundant protein database resulting in 1,091 putative identifications that were assigned biological roles using a classification system adapted from Riley¹⁵ (Table 2). The 1,590 predicted genes had an average size of 945 bp, similar to that observed in other prokaryotes^{7–9}, and no genome-wide strand bias was observed (Fig. 2). More than 70% of the predicted proteins in *H. pylori* have a calculated isoelectric point (pI) greater than 7.0, compared to ~40% in *H. influenzae* and *E. coli*. The basic amino acids, arginine and lysine, occur twice as frequently in *H. pylori* proteins as in those of *H. influenzae* and *E. coli*, perhaps reflecting an adaptation of *H. pylori* to gastric acidity.

Paralogous families. Ninety-five paralogous gene families comprising 266 gene products (16% of the total) were identified (www.tigr.org/tdb/mdb/hpdbh/hpdbh.html). Of these, 67 (173 proteins) have an assigned role. Sixty-four have only 2 members, while the porin/adhesin-like outer membrane protein family (Fig. 2) is the largest with 32 members. The largest number of paralogues with assigned roles fall into the functional categories of cell

envelope, transport and binding proteins, and proteins involved in replication. The large number of cell envelope proteins might reflect either a reduced biosynthetic capacity or a need to adapt to the challenging gastric environment.

Cell division and protein secretion

The gene content of *H. pylori* suggests that the basic mechanisms of replication, cell division and secretion are similar to those of *E. coli* and *H. influenzae*. However, important differences are noted. For example, apparently missing from the *H. pylori* genome are orthologues of DnaC, MinC, and the secretory chaperonin, SecB. In oriC-type primosome formation, the DnaB and DnaC proteins form a B–C complex that delivers the DnaB helicase to the developing primosome complex¹⁶. The apparent absence of DnaC in *H. pylori* suggests that either a novel mechanism for recruiting DnaB exists or a DnaC orthologue with no detectable sequence similarity is present. Similar arguments can be made for other seemingly missing important functions.

H. pylori has a classical set of bacterial chaperones (DnaK, DnaJ, CbpA, GrpE, GroEL, GroES, and HtpG). The transcriptional regulation of *H. pylori* chaperone genes is likely to be different from that in *E. coli*, as it seems not to have the sigma factors that upregulate chaperone synthesis in *E. coli* (heat-shock sigma 32 and stationary-phase sigma S).

In addition to the SecA-dependent secretory pathway, *H. pylori* has two specialized export systems. One is associated with the *cag* pathogenicity island¹¹ and the other is the flagellar export pathway which is assembled from orthologues of FliH, FliI, FliP, FliB, FliQ, FliR and FliP¹⁷. Apparently absent from *H. pylori* is a type IV signal peptidase and orthologues of the dsbABC system, which in other species are required for the maturation of pili and pilin-like structures¹⁸ and assembly of surface structures involved in virulence and DNA transformation¹⁹.

Recombination, repair and restriction systems

Systems for homologous recombination and post-replication, mismatch, excision and transcription-coupled repair appear to be present in *H. pylori*. Also present are genes with similarity to DNA glycosylases which have associated AP endonuclease activity. The RecBCD pathway, which mediates homologous recombination and double-strand break repair, and RecT and RecE orthologues, proteins involved in strand exchange during recombination²⁰, seem to be absent. The ability of *H. pylori* to perform mismatch repair is suggested by the presence of methyl transferases, mutS and uvrD. However, orthologues of MutH and MutL were not identified. Components of an SOS system also appear to be absent.

Bacteria commonly use restriction and modification systems to degrade foreign DNA. In *H. pylori*, this defence system is well developed with eleven restriction-modification systems identified on the basis of gene order and similarity to endonucleases, methyltransferases, and specificity subunits. Three type I, one type II, and three type III systems were identified, as well as four type III systems, including the recently identified epithelial responsive

Figure 1 Linear representation of the *H. pylori* 26695 chromosome illustrating the location of each predicted protein-coding region, RNA gene, and repeat elements in the genome. Symbols are as follows: ++, Co²⁺, Zn²⁺, Cd²⁺; ?, unknown; A/G/S, D-alanine/glycine/D-serine; B12, B12/ferric siderophores; E, glutamate; Mo, molybdenum; P, proline; P/G, proline/glycine betaine; Q, glutamine; S, serine; a-k, α-ketoglutarate; a/o, arginine/ornithine; aa, amino acids (specificity unknown); aa2, dipeptides; aaX, oligopeptides; fum, fumarate, succinate; glu, glucose/galactose; h, hemin; lac, L-lactate; mal, malate 2-oxoglutarate; nic, nicotinamide mononucleotides; pyr, pyrimidine nucleosides. Numbers associated with tRNA symbols represent the number of tRNAs at a locus. Numbers associated with GES represent the number of membrane-spanning domains according to the Goldman, Engelman and Steitz scale as calculated by TopPred⁴⁷.

endonuclease, *iceA1*, and its associated DNA adenine methyltransferase (M. HypI) genes^{21,22}. In addition to the complete systems, seven adenine-specific, and four cytosine-specific methyltransferases, and one of unknown specificity were found. Each of these has an adjacent gene with no database match, suggesting that they may function as part of restriction-modification systems.

Transcription and translation

Although analysis of gene content suggests that *H. pylori* has a basic transcriptional and translational machinery similar to that of *E. coli*, interesting differences are observed. For example, no genes for a catalytic activity in tRNA maturation (*rnd*, *rph*, or *rnpB*) were identified and of the three known ribonucleases involved in mRNA degradation, only polyribonucleotide phosphorylase was found. Twenty-one genes coding for 18 of the 20 tRNA synthetases normally required for protein biosynthesis were found.

As in most other completely sequenced bacterial genomes, the gene for glutamyl-tRNA synthetase, *glnS*, is missing, and the existence of a transamidation process is assumed. It is also possible that the product of the second glutamyl-tRNA synthetase gene, *gltX*, present in *H. pylori*, may have acquired the glutamyl-tRNA synthetase function. *H. pylori* provides the first example of a bacterial genome apparently lacking an asparaginyl-tRNA synthetase gene, *asnS*. A transamidation process to form *Asn-tRNA^{Asn}* from *Asp-tRNA^{Asn}* has been reported for the archaeon *Haloferax volcanii*²² and may also operate in *H. pylori*. Most intriguing, however, is the finding that in *H. pylori* the genes encoding the β and β' subunits of RNA polymerase are fused. In all studied prokaryotes the two genes are contiguous, but separate, and are part of the same transcriptional unit. Whether this gene fusion in *H. pylori* results in a fused protein, or whether the transcriptional or translational product of the fusion is subject to splicing, is currently not known. It is worth noting that an artificial fusion of the *E. coli*

rpoB and *rpoC* genes is viable and results in a transcriptional complex, which has the same stoichiometry as the native complex (K. Severinov, personal communication).

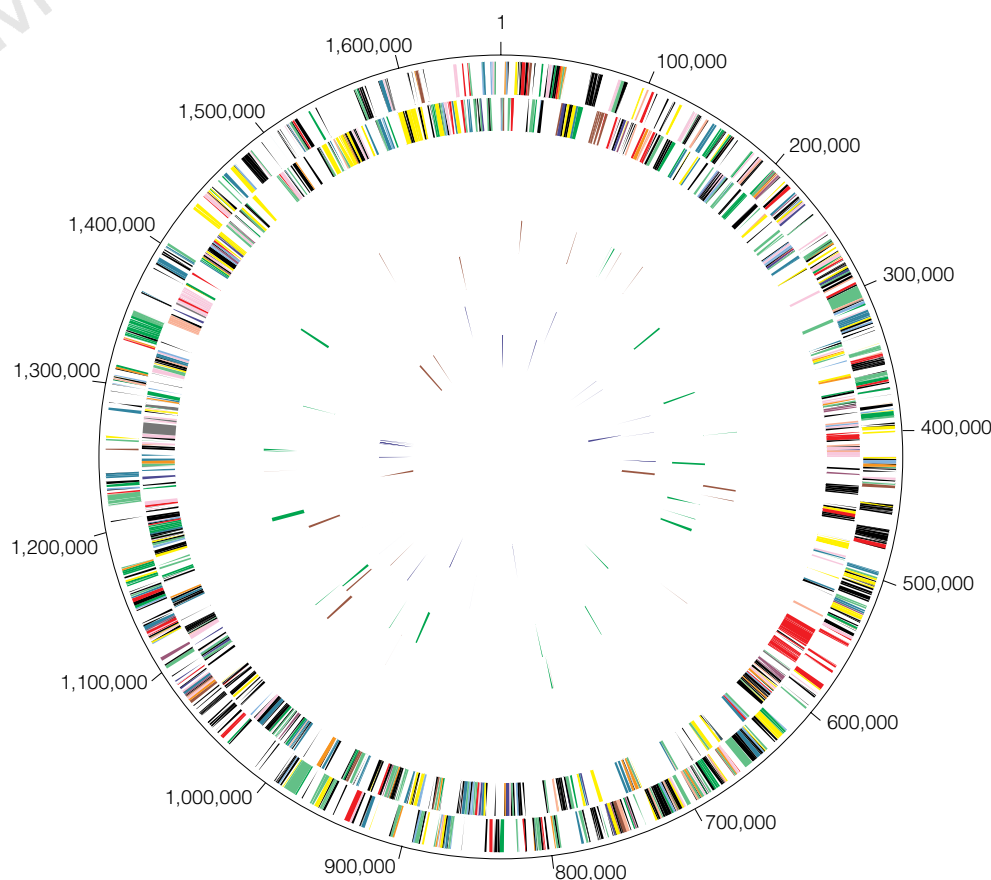
Adhesion and adaptive antigenic variation

Most pathogens show tropism to specific tissues or cell types and often use several adherence mechanisms for successful attachment. *H. pylori* may use at least five different adhesins to attach to gastric epithelial cells⁵. One of them, HpaA (HP0797), was previously identified as a lipoprotein in the flagellar sheath and outer membrane^{5,23}. In addition to the HpaA orthologue, we have identified 19 other lipoproteins. Few have an identifiable function, but some are likely to contribute to the adherence capacity of the organism.

Two adhesins^{24–26}, one of which mediates attachment to the Lewis^b histo-blood group antigens, belong to the large family of outer membrane proteins (OMP) (Fig. 3) (T. Boren and R. Haas, personal communication). It is conceivable that other members of these closely related proteins also act as adhesins. Given the large number of sequence-related genes encoding putative surface-exposed proteins, the potential exists for recombinational events leading to mosaic organization. This could be the basis for antigenic variation in *H. pylori* and an effective mechanism for host defence evasion, as seen in *M. genitalium*²⁷.

At least one other mechanism for antigenic variation could operate in *H. pylori*. The DNA sequence at the beginning of eight genes, including five members of the OMP family, contain stretches of CT or AG dinucleotide repeats (Table 3a). In addition, poly(C) or poly(G) tracts occur within the coding sequence of nine other genes (Table 3b). Slipped-strand mispairing within such repeats are documented features of one mechanism of genotypic variation^{28,29}. These mechanisms may have evolved in bacterial pathogens to increase the frequency of phenotypic variation in genes involved in

Figure 2 Circular representation of the *H. pylori* 26695 chromosome. Outer concentric circle: predicted coding regions on the plus strand classified as to role according to the colour code in Fig. 1 (except for unknowns and hypotheticals, which are in black). Second concentric circle: predicted coding regions on the minus strand. Third and fourth concentric circles: IS elements (red) and other repeats (green) on the plus and minus strand, respectively. Fifth and sixth concentric circles: tRNAs (blue), rRNAs (red), and sRNAs (green) on the plus and minus strand, respectively.



N-Terminal peptide sequences:
 HopA EE—DGFVSVQYQIE—AQVQV
 HopB EE—DGFVSVQYQIE—AQVQV
 HopC ED—DGFVSVQYQIE—AQVQV
 HopD ED—DGFVSVQYQIE—AQVQV
 HopE EE—DGFVSVQYQIE—AQVQV

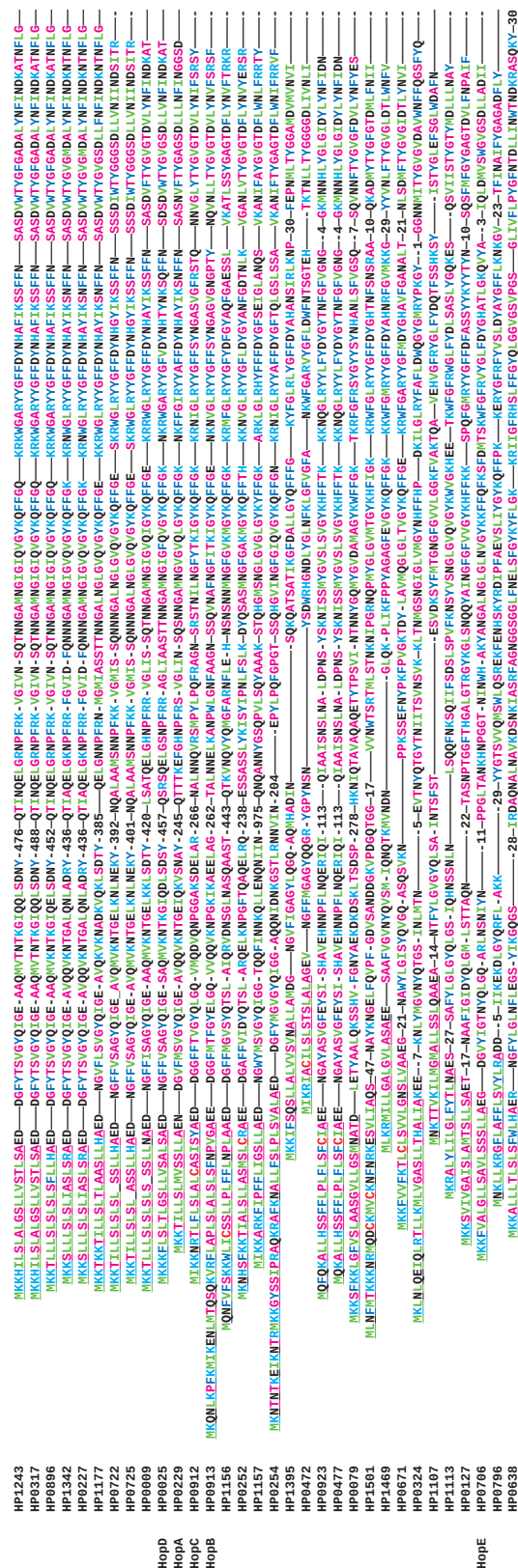
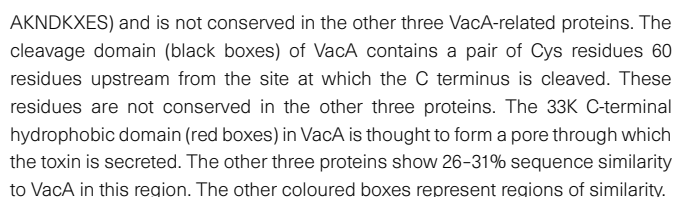
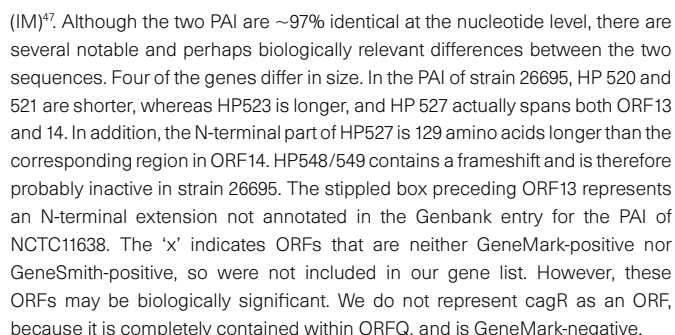


Figure 3 Multiple sequence alignment of members of the outer membrane protein family of *H. pylori*. These proteins were identified as OMPs based on the characteristic alternating hydrophobic residues at their carboxy termini. All members of this family have one domain of similarity at the amino-terminal end and seven domains of similarity at their carboxy-terminal end. Note that the first 11 of these OMPs share extensive similarity over their entire length. Four of the OMPs were identified as porins (Hops) based on identity to published amino-terminal sequences, represented at the top of the alignment⁵⁰. The most likely

candidate for HopD is HP0913, which has 15 matches to the first 20-residue N-terminal peptide sequence⁵⁰. These differences may be due to strain variability. The program Signal-P⁴⁸ was used to identify cleavage sites and signal peptides (underlined). Four of the OMPs have TTG start codons (HP1156, HP0252, HP1113, HP0796). Numbers embedded in the sequences represent amino acids omitted from the alignment. The star symbols indicate that HP722, HP725 and HP9 proteins contain a frameshift in their signal-peptide-coding region. These frameshifts are associated with the presence of dinucleotide repeats (Table 3).

VacA induces the formation of acidic vacuoles in host epithelial cells, and its presence is associated epidemiologically with tissue damage and disease³¹. VacA may not be the only ulcer-causing factor as 40% of *H. pylori* strains do not produce detectable amounts of the cytotoxin *in vitro*⁵. Sequence differences at the amino terminus and central regions are noted among VacA proteins derived from Tox⁺ and Tox⁻ strains³¹. This Tox⁺ *H. pylori* strain contains the more toxigenic SlA/m1 type cytotoxin and three additional large proteins with moderate similarities to the carboxy-terminal end of the active



As with other pathogens, *H. pylori* probably requires an iron-scavenging system for survival in the host⁵. Genome analysis suggests that *H. pylori* has several systems for iron uptake. One is analogous to the siderophore-mediated iron-uptake *fec* system of *E. coli*³⁴, except that it lacks the two regulatory proteins (FecR and FecI) and is not organized in a single operon. Unlike other studied systems, *H. pylori* has three copies of each of *fecA*, *exbB* and *exbD*. A second system, consisting of a *feoB*-like gene without *feoA*, suggests that *H. pylori* can assimilate ferrous iron in a fashion similar to the anaerobic *feo* system of *E. coli*. Other systems for iron uptake present in *H. pylori* consist of the three *frpB* genes which encode proteins similar to either haem- or lactoferrin-binding proteins. Finally, *H. pylori* contains NapA, a bacteriostatin³⁴, and Pfr, a non-haem cytoplasmic iron-containing protein used for storage of iron³⁵. The global ferric uptake regulator (Fur) characterized in other bacteria is also present in *H. pylori*. Consensus

H. pylori motility is essential for colonization³⁶. It enables the bacterium to spread into the viscous mucous layer covering the gastric epithelium. At least forty proteins in the *H. pylori* genome appear to be involved in the regulation, secretion and assembly of the flagellar architecture. As has been reported for the *flaA* and *flaB* genes, we identified sigma 28 and sigma 54-like promoter elements upstream of many flagellar genes, underscoring the complexity of the transcriptional regulation of the flagellar regulon⁵.

H. pylori is unusual among pathogenic bacteria in its ability to colonize host cells in an environment of high acidity. As it enters the gastric environment by oral ingestion, the organism is transiently subjected to the extreme pH of the lumen side of the gastric mucous layer (pH ~2). The survival of *H. pylori* in acidic environments is probably due to its ability to establish a positive inside-membrane potential³⁷ and subsequently to modify its microenvironment through the action of urease and the release of factors that inhibit acid production by parietal cells⁵. A switch in membrane polarity provides an electrical barrier that prevents the entry of protons (H^+). A positive cell interior can be created by the active extrusion of anions or by a proton diffusion potential. The latter model appears more likely as no clear mechanism for electrogenic anion efflux is apparent in the genome. A proton diffusion potential would require the anion permeability of the cytoplasmic membrane to be low and, thus far, only three anion transporters have been identified. However, it remains to be determined whether anion conductances are associated with other proteins: the MDR-like transporters (HP600, HP1082 and HP1206) or hypotheticals. Although it has been suggested that proton-translocating P-type ATPases could mediate survival in acid conditions by the extrusion of protons from the cytoplasm³⁸, this idea is not supported by the identified transporter

HP no.	ID	No. of repeats	Gene status	Poly(A) or Poly(T) tracts in 5' intergenic region
--------	----	----------------	-------------	---

Nucleotide sequence at the beginning of HP0722 showing the CT dinucleotide repeat and the poly T tract. The putative ribosome binding site is shown in green. Translation starting at the designated methionine leads to a truncated product. The addition or deletion of two CT repeats, by 'slipped-strand mispairing', will restore the frame.

(b) Homopolymeric poly(C) and poly(G) tracts within coding sequence

HP no.	ID	Tract length	Gene status
58	Hypo	C15	Off
217	Hypo	G12	On
379	fucosyl transf.	C13	On
464	Typel R	C15	On
619	glycos. transf.	C13	Truncated
651	Hypo	C13	On
1353	Hypo	C15	Truncated
1471	TypelIS-R	G14	On
1522	Methyl ase	G12	Truncated

Genes possibly regulated by homopolymeric poly(A) or poly(T) tracts in 5' intergenic regions

HP no.	ID	Tract	HP no.	ID	Tract	HP no.	ID	Tract
9	OMP	A14	25	OMP	T15	208	<i>rfaJ</i>	A11
227	OMP	T14	228	IMP	A14	349	<i>pyrG</i>	T15
350	IMP	A15	547	<i>cagA</i>	A14	629	Hypo	T15
722	OMP	T16	725	OMP	T14	733	Hypo	T13
876	<i>frpB</i>	T16	896	OMP	A14	912	OMP	T13
1342	OMP	A14	1400	<i>fecA</i>	A16			

genes. The P-type ATPase sequences in *H. pylori* (*copAP*, HP791, and HP1503) are more closely related to divalent cation transporters than to ATPases with specificity for protons or monovalent cations. One of them, HP0791, is involved in Ni^{2+} supply, an essential component of urease activity³⁹. The others may be involved in the elimination of toxic metals from the cytoplasm and not in pH regulation.

Additional mechanisms of pH homeostasis may well contribute to *H. pylori* survival. A change in protein content observed in response to a shift of extracellular pH from 7.5 to 3.0 suggests the presence of an acid-inducible response⁴⁰. Although *H. pylori* lacks most orthologues of the genes that are acid-induced in *E. coli* and *Salmonella typhimurium*, including the amino-acid decarboxylases and formate hydrogen lyase, certain virulence factors, outer membrane

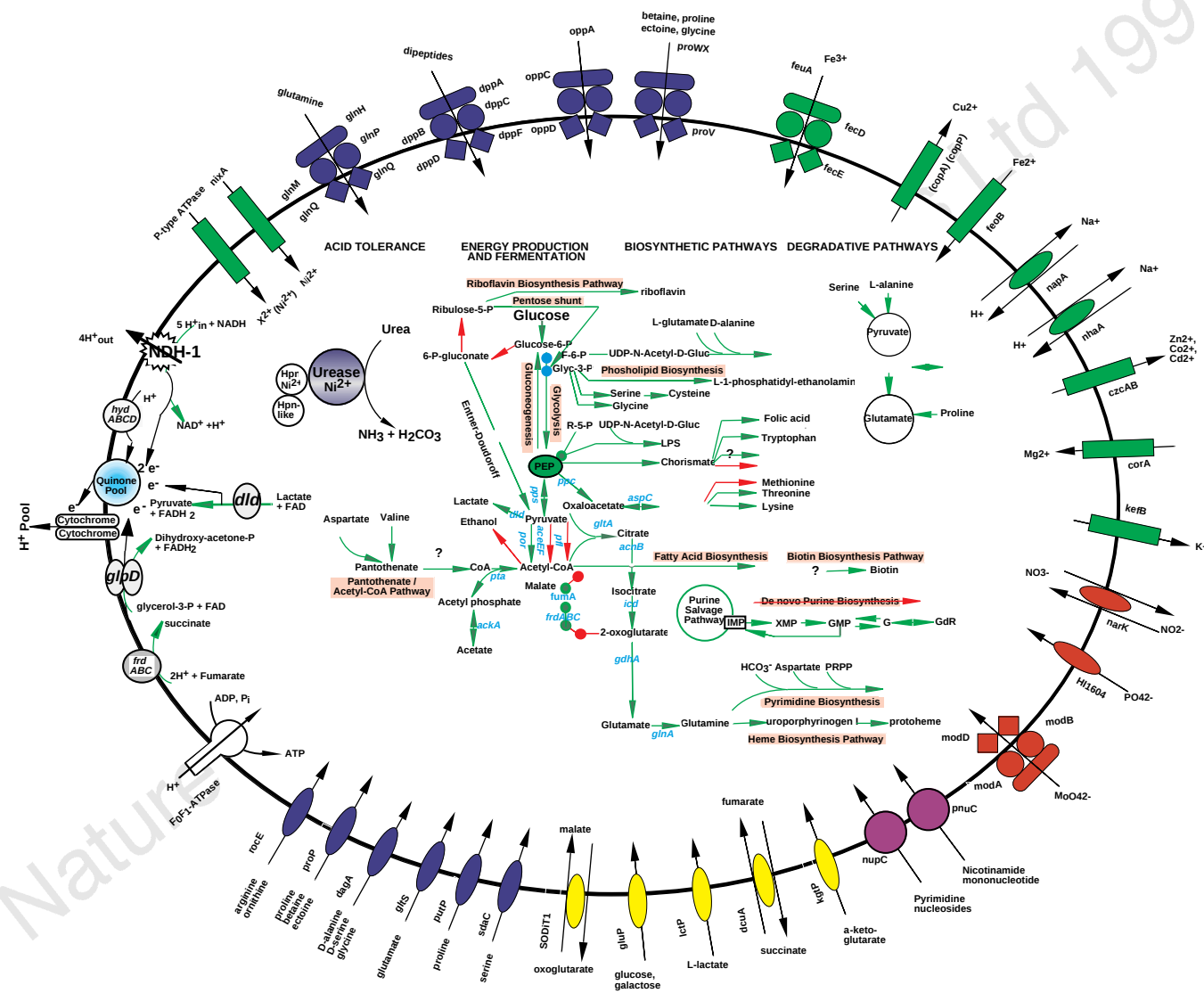


Figure 6 Solute transport and metabolic pathways of *Helicobacter pylori*. Transporters identified by sequence comparisons are characteristic of Gram-negative bacteria. Colours correspond to transport role categories defined by Riley¹⁵: blue, amino acids, peptides and amines; red, anions; yellow, carbohydrates, organic alcohols and acids; green, cations; and purple, nucleosides, purines and pyrimidines. Numerous permeases (ovals) with specificity for amino acids (*recE*, *proP*, *dagA*, *gltS*, *putP* and *sdaC*) or carbohydrates (*SODiTi*, *gluP*, *lactP*, *cduA*, *kgtP*) import organic nutrients. Structurally related permease proteins maintain ionic homeostasis by transporting HPO_4^{2-} (*H1604*), NO_3^- (*narK*), and Na^+ (*nhaA*, *napA*). Primary active-transport systems, independent of the proton cycle, are also apparent. Included in this group are ATP-binding protein-cassette (ABC) transporters (composite figures of 2 diamonds, 2 circles, 1 oval) for the uptake of oligopeptides (*oppACD*), dipeptides (*dppABCD*), proline (*proWX*), glutamine (*glnHMPQ*), molybdenum (*modABD*), and iron III (*fecED*), P-type ATPases that extrude toxic metals from the cell (*copAP* and *cadA*), and the glutathione-regulated potassium-efflux protein (*kefB*). Transporters for the accumulation of ionic cofactors are encoded by *nixA* (Ni^{2+} for urease activation), *corA* (Mg^{2+} for phosphohydrolases, phosphotransferases, ATPases) and *feoB* (Fe^{2+}

import under anaerobic conditions for cytochromes, catalase). An integrated view of the main components of the central metabolism of *H. pylori* strain 26695 is presented. The use of glucose as the sole carbohydrate source is emphasized. Urease, a multisubunit Ni^{2+} -binding enzyme, is crucial for colonization and for survival of *H. pylori* at acid pH, and is indicated as a complex (purple circle) with Hpn, a Ni^{2+} -binding cofactor, and a newly identified Hpn-like protein (HP1432). A question mark is attached to pathways that could not be completely elucidated. Pathways or steps for which no enzymes were identified are represented by a red arrow. Pathways for macromolecular biosynthesis (RNA, DNA and fatty acids) have been omitted. *ackA*, acetate kinase; *acnB*, aconitase B; *aspC*, aspartate aminotransferase; *dld*, D-lactate dehydrogenase; *gdhA*, glutamate dehydrogenase; *glnA*, glutamine synthetase; *gltA*, citrate synthase; *HydABC*, hydrogenase complex; *icd*, isocitrate dehydrogenase; *pfl*, pyruvate formate lyase; *por*, pyruvate ferredoxin oxidoreductase; *ppc*, phosphoenolpyruvate carboxylase; *pps*, phosphoenolpyruvate synthase; *pta*, phosphate acetyltransferase; *gldD*, glycerol-3-phosphate dehydrogenase; NDH-1, NADH-ubiquinone oxidoreductase complex.

proteins, sensor-regulator pairs and other proteins may be acid-induced.

Regulation of gene expression

Bacteria regulate the transcription of their genes in response to many environmental stimuli, such as nutrient availability, cell density, pH, contact with target tissue, DNA-damaging agents, temperature and osmolarity. In the case of pathogens, the regulated expression of certain key genes is essential for successful evasion of host responses and colonization, adaptation to different body sites, and survival as the pathogen passes to new hosts. In *H. pylori*, global regulatory proteins are less abundant than in *E. coli*. For example, orthologues of many DNA-binding proteins that regulate the expression of certain operons such as OxyR (oxidative stress), Crp (carbon utilization), RpoH (heat shock), and Fnr (fumarate and nitrate regulation) are absent. Only four *H. pylori* proteins have a perfect match to helix–turn–helix (HTH) motifs, a signature of transcription factors; a putative heat-shock protein (HspR), two proteins with no database match (HP1124 and HP1349) and SecA, a component of the general secretory machinery. In contrast, 34 proteins containing an HTH motif were found in *H. influenzae* and 148 in *E. coli*. We identified several other putative regulatory functions, including SpoT and CstA for 'stringent response' to amino-acid starvation and to carbon starvation, respectively.

Environmental response requires sensing changes and transmission of this information to cellular regulatory networks. Two-component regulator systems, consisting of a membrane histidine kinase sensor protein and a cytoplasmic DNA-binding response regulator, provide a well studied mechanism for such signal transduction. Four sensor proteins and seven response regulators were found in *H. pylori*, similar to the number found in *H. influenzae*⁷. This is approximately one third the number found in *E. coli* which, in contrast to *H. pylori* and *H. influenzae*, may be exposed to more environments.

Metabolism

Metabolic pathway analysis of the *H. pylori* genome suggests the following features. *H. pylori* uses glucose as the only source of carbohydrate and the main source for substrate-level phosphorylation. It also derives energy from the degradation of serine, alanine, aspartate and proline. The glycolysis–gluconeogenesis metabolic axis constitutes the backbone of energy production and the start point of many biosynthetic pathways. The biosynthesis of peptidoglycan, phospholipids, aromatic amino acids, fatty acids and cofactors is derived from acetyl-CoA or from intermediates in the glycolytic pathway (Fig. 6). The metabolism of pyruvate reflects the microaerophilic character of this organism. Neither the aerobic pyruvate dehydrogenase (*aceEF*) nor the strictly anaerobic pyruvate formate lyase (*pfl*) associated with mixed-acid fermentation are present. The conversion of pyruvate to acetyl CoA is performed by the pyruvate ferredoxin oxidoreductase (POR), a four-subunit enzyme thus far only described in hyperthermophilic organisms⁴¹. The tricarboxylic acid cycle (TCA) is incomplete and the glyoxylate shunt is absent. The analysis of degradative pathways, uptake systems and biosynthetic pathways for pyrimidine, purine and haem suggests that *H. pylori* uses several substrates as nitrogen source, including urea, ammonia, alanine, serine and glutamine. The assimilation of ammonia, an abundant product of urease activity, is achieved by the glutamine synthase enzyme and α -ketoglutarate is transformed into glutamate by glutamate dehydrogenase rather than by the glutamate synthase enzyme.

In *H. pylori*, proton translocation is mediated by the NDH-1 dehydrogenase and the different cytochromes, including the primitive-type cytochrome cbb3 (Table 2). Four respiratory electron-generating dehydrogenases have been identified, glycerol-3-phosphate dehydrogenase (GlpD), D-lactate dehydrogenase, NADH–ubiquinone oxidoreductase complex (NDH-1), and a hydrogenase complex (HydABC). Our analysis also suggests that

H. pylori is not able to use nitrate, nitrite, dimethylsulphoxide, trimethylamine N-oxide or thiosulphate as electron acceptors. Much of our metabolic analysis is supported by experimental evidence^{41,42}.

Evolutionary relationships of *H. pylori*

H. pylori is currently classified in the Proteobacteria, a large, diverse division of Gram-negative bacteria which includes two other completely sequenced species, *H. influenzae* and *E. coli*. Given this taxonomic placement, based primarily on 16S rRNA sequence comparisons, one might expect the proteins of *H. pylori* more closely to resemble their *H. influenzae* and *E. coli* homologues rather than those in other genomes such as *Synechocystis* sp., *M. genitalium*, *M. pneumoniae*, *M. jannaschii*, and *Saccharomyces cerevisiae*. This is indeed the case for many proteins. There are, however, many examples of *H. pylori* proteins in amino-acid biosynthesis, energy metabolism, translation and cellular processes that have greater sequence similarity to those found in non-Proteobacteria. For example, Dhs1, the initial enzyme in the chorismate biosynthesis pathway is 75.5% similar to *Arabidopsis thaliana* chloroplast Dhs1 gene product, and has minimal sequence similarity to the equivalent *E. coli* AroH, AroF or AroG gene products. The remaining enzymes in this pathway have strong sequence similarity to their *E. coli* counterpart. Similarly, the *H. pylori* prephenate dehydrogenase (TyrA), which converts chorismate to tyrosine, and six out of 15 enzymes in the aspartate amino acid biosynthetic pathways, resemble those from *B. subtilis*. A similar pattern can be seen in a different functional category. Nearly all *H. pylori* tRNA synthetases have eubacterial homologues, mostly with best matches to Proteobacteria species. However, histidyl-tRNA synthetase shows several amino-acid sequence signatures in common with eukaryotic and archaeal (*M. jannaschii*) homologues.

Such observations of discordant sequence similarity are often interpreted as evidence of lateral gene transfer in the evolutionary history of an organism. It is also possible that *H. pylori* diverged early from the lineage that led to the gamma Proteobacteria, and retained more ancient forms of enzymes that have been subsequently replaced or have diverged extensively in *H. influenzae* and *E. coli*.

Conclusion

Our whole-genome analysis of *H. pylori* gives new insight into its pathogenesis, acid tolerance, antigenic variation and microaerophilic character. The availability of the complete genome sequence will allow further assessment of *H. pylori* genetic diversity. This is an important aspect of *H. pylori* epidemiology as allelic polymorphism within several loci has already been associated with disease outcome^{5,21,31}. The extent of molecular mimicry between *H. pylori* and its human host, an underappreciated topic, can now be fully explored⁴³. The identification of many new putative virulence determinants should allow critical tests of their roles and thus new insight into mechanisms of initial colonization, persistence of this bacterium during long-term carriage, and the mechanisms by which it promotes various gastroduodenal diseases.

Methods

H. pylori strain 26695 (ref. 44) was originally isolated from a patient in the United Kingdom with gastritis (K. Eaton, personal communication) and was chosen because it colonizes piglets and elicits immune and inflammatory responses. It is also toxigenic, and transformable, and thus amenable to mutational tests of gene function.

The *H. pylori* genome sequence was obtained by a whole-genome random sequencing method previously applied to genomes of *Haemophilus influenzae*⁷, *Mycoplasma genitalium*⁸, and *Methanococcus jannaschii*⁹. Ninety-two per cent of the genome was covered by at least one λ clone and only 0.56% of the genome had single-fold coverage.

Open reading frames (ORFs) and predicted coding regions were identified using three methods. The predicted protein-coding regions were initially defined by searching for ORFs longer than 80 codons. Coding potential analysis of the entire genome was performed with a version of GeneMark⁴⁵ trained with a set of *H. pylori* ORFs longer than 600 nucleotides. Coding sequences and potential starts of translation were also determined using GeneSmith (H.S., unpublished), a program that evaluates ORF length, separation of ORFs and overlap and quality of ribosome binding site. ORFs with low GeneMark coding potential, no database match, and not retained by GeneSmith were eliminated. GeneSmith identified 25 ORFs that are smaller than 100 codons, had no database match and were GeneMark negative. Frameshifts were detected by inspecting pairwise alignments, families of orthologues (similar proteins derived from different species) and paralogues (similar proteins from within the same organism), and regions containing homopolymer stretches and dinucleotide repeats. Ambiguities were resolved by an alternative sequencing chemistry (terminator reactions), and by sequencing PCR products obtained using the genomic DNA as template. Frameshifts that remain in the genome are considered authentic and not sequencing artefacts.

To determine their identity, ORFs were searched against a non-redundant amino-acid database as previously described⁹. ORFs were also analysed using 175 hidden Markov models constructed for a number of conserved protein families (pfam v1.0) using hmmer⁴³. In addition, all ORFs were searched against the prosite motif database using MacPattern⁴⁶. Families of paralogues were constructed by pairwise searches of proteins using FASTA. Matches that spanned at least 60% of the smaller of the protein pair were retained and visually inspected.

A unix version of the program TopPred⁴⁷ was used to identify membrane-spanning domains (MSD) in proteins. Six hundred and sixty three proteins containing at least one MSD were found; of these, 300 had 2 potential MSDs or more. The presence of signal peptides and the probable position of the cleavage site in secreted proteins were detected using Signal-P, a neural net program that had been trained on a curated set of secreted proteins from Gram-negative bacteria⁴⁸. 367 proteins were predicted to have a signal peptide. Lipoproteins were identified by scanning for the presence of a lipobox in the first 30 amino acids of every protein; 20 lipoproteins were identified, eighteen of which were Signal-P positive. Outer-membrane proteins were found by searching for aromatic amino acids at the end of the proteins.

Homopolymer and dinucleotide repeats were found by using RepScan (H.O.S., unpublished) which finds direct repeats of any length. All features identified using these programs were validated by visual inspection to remove false positives. Metabolic pathways were curated by hand and by reference to EcoCyc⁴⁹.

Received 16 May; accepted 1 July 1997.

- Warren, J. R. & Marshall, B. Unidentified curved bacilli on gastric epithelium in active chronic gastritis. *Lancet* **1**, 1273–1275 (1983).
- Cover, T. L. & Blaser, M. J. *Helicobacter pylori* infection, a paradigm for chronic mucosal inflammation: pathogenesis and implications for eradication and prevention. *Adv. Int. Med.* **41**, 85–117 (1996).
- Mobley, H. L. T., Island, M. D. & Hausinger, R. P. Molecular Biology of Microbial Ureases. *Microbiol. Rev.* **59**, 451–480 (1995).
- Go, M. F. & Graham, D. Y. How does *Helicobacter pylori* cause duodenal ulcer disease: The bug, the host, or both? *J. Gastroenterol. Hepatol.* (suppl.) **9**, 8–12 (1994).
- Labigne, A. & de Reuse, H. Determinants of *Helicobacter pylori* pathogenicity. *Infect. Agents Disease* **5**, 191–202 (1996).
- Clemens, J. et al. Impact of infection by *Helicobacter pylori* on the risk and severity of endemic cholera. *J. Inf. Dis.* **171**, 1653–1656 (1995).
- Fleischmann, R. D. et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512 (1995).
- Fraser, C. M. et al. The *Mycoplasma genitalium* genome sequence reveals a minimal gene complement. *Science* **270**, 397–403 (1995).
- Bult, C. J. et al. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058–1073 (1996).
- Winans, S. C., Burns, D. L. & Christie, P. J. Adaptation of a conjugal transfer system for the export of pathogenic macromolecules. *Trends Microbiol.* **4**, 64–68 (1996).
- Censini, S. et al. Cag, a pathogenicity island of *Helicobacter pylori*, encodes type-I-specific and disease-associated virulence factors. *Proc. Natl Acad. Sci. USA* **93**, 14648–14653 (1996).
- <http://genome.wustl.edu/eddy/low/rNAScan-SE-Manual/Manual.html>
- Akopyants, N. S., Kersulyte, D. & Berg, D. E. DNA rearrangement in the 40 kb cag (virulence) region in the *Helicobacter pylori* genome. *Gut* **39** (suppl. 2), A67 (1996).
- Marczynski, G. T. & Shapiro, L. Bacterial chromosome origins of replication. *Curr. Opin. Gen. Dev.* **3**, 775–782 (1993).
- Riley, M. Functions of gene products of *Escherichia coli*. *Microbiol. Rev.* **57**, 862–952 (1993).
- Kornberg, A. & Baker, T. A. Replication mechanisms and operations in DNA replication. (ed. Kornberg, A. & Baker, T.) 471–510 (Freeman, New York, 1992).

- Macnab, R. M. in *Escherichia coli and Salmonella Cellular and Molecular Biology* (eds Neidhardt, F. C. et al.) 123–145 (ASM, Washington DC, 1996).
- Strom, M. S., Nunn, D. N. & Lory, S. Posttranslational processing of type IV prepilin and homologs by PilD of *Pseudomonas aeruginosa*. *Meth. Enzymol.* **235**, 527–540 (1994).
- Bardwell, J. C. Building bridges: disulphide bond formation in the cell. *Mol. Microbiol.* **14**, 199–205 (1994).
- Linn, S. in *Escherichia coli and Salmonella Cellular and Molecular Biology* (eds Neidhardt, F. C. et al.) 764–772 (ASM, Washington D.C., 1996).
- Peek, R. M., Thompson, S. A., Atherton, J. C., Blaser, M. J. & Miller, G. G. Expression of iceA, a novel ulcer-associated *Helicobacter pylori* gene, is induced by contact with gastric epithelial cells and is associated with enhanced mucosal IL-8. *Gut* **39** (suppl. 2), A71 (1996).
- Curnow, A. W., Ibba, M. & Soll, D. tRNA-dependent asparagine formation. *Nature* **382**, 589–590 (1996).
- Jones, A. C., Foynes, S., Cockayne, A. & Penn, C. W. Gene cloning of a flagellar sheath protein of *Helicobacter pylori* shows its identity with the putative adhesin, HpaA. *Gut* **39** (suppl. 2), A62 (1996).
- Boren, T., Falk, P., Roth, K. A., Larson, G. & Normark, S. Attachment of *Helicobacter pylori* to human gastric epithelium mediated by blood group antigens. *Science* **262**, 1892–1895 (1993).
- Ilyer, D. et al. The *Helicobacter pylori* blood group antigen binding adhesin. *Gut* **39** (suppl. 2), A55 (1996).
- Odenbreit, S., Till, M. & Haas, R. Optimized blaM-transposon shuttle mutagenesis of *Helicobacter pylori* allows identification of novel genetic loci involved in bacterial virulence. *Mol. Microbiol.* **20**, 361–373 (1996).
- Peterson, S. N. et al. Characterization of repetitive DNA in the *Mycoplasma genitalium* genome: possible role in the generation of antigenic variation. *Proc. Natl Acad. Sci. USA* **92**, 11829–11833 (1995).
- Moxon, E. R., Rainey, P. B., Nowak, M. A. & Lenski, R. E. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr. Biol.* **4**, 24–33 (1994).
- Jonsson, A. B., Nyberg, G. & Normark, S. Phase variation of gonococcal pili by frameshift mutation in pilC, a novel gene for pilus assembly. *EMBO J.* **10**, 477–488 (1991).
- Tummuru, M. K. R., Sharma, S. A. & Blaser, M. J. *Helicobacter pylori* picB, a homologue of the *Bordetella pertussis* toxin secretion protein, is required for induction of IL-8 in gastric epithelial cells. *Mol. Microbiol.* **18**, 867–876 (1995).
- Atherton, J. C. et al. Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. Association of specific vacA types with cytotoxin production and peptic ulceration. *J. Biol. Chem.* **270**, 17771–17777 (1995).
- Moran, A. P. The role of lipopolysaccharide in *Helicobacter pylori* pathogenesis. *Aliment. Pharmacol. Ther.* **10** (suppl. 1), 39–50 (1996).
- Baker, P. J. et al. Molecular structures that influence the immunomodulatory properties of the lipid A and inner core region oligosaccharides of bacterial lipopolysaccharides. *Infect. Immun.* **62**, 2257–2269 (1994).
- Earhart, C. F. in *Escherichia coli and Salmonella Cellular and Molecular Biology* (eds Neidhardt, F. C. et al.) 1075–1090 (ASM, Washington DC, 1996).
- Evans, D. J., Jr, Evans, D. G., Lampert, H. C. & Nakano, H. Identification of four new prokaryotic bacterioferritins, from *Helicobacter pylori*, *Anabaena variabilis*, *Bacillus subtilis* and *Treponema pallidum*, by analysis of gene sequences. *Gene* **153**, 123–127 (1995); Frazier, B. A. et al. Paracrystalline inclusions of a novel ferritin containing nonheme iron, produced by the human gastric pathogen *Helicobacter pylori*: evidence for a third class of ferritins. *J. Bacteriol.* **175**, 966–972 (1993).
- Suerbaum, S. The complex flagella of gastric *Helicobacter* species. *Trends Microbiol.* **3**, 168–170 (1995).
- Martin, A., Zychlinsky, E., Keyhan, M. & Sachs, G. Capacity of *Helicobacter pylori* to generate ionic gradients at low pH is similar to that of bacteria which grow under strongly acidic conditions. *Infect. Immun.* **64**, 1434–1436 (1996).
- Melchers, K. et al. Cloning and membrane topology of a P type ATPase from *Helicobacter pylori*. *J. Biol. Chem.* **271**, 446–457 (1996).
- Melchers, K. et al. Cloning and analysis of two P type ion pumps of *Helicobacter pylori*, a cation resistance ATPase and a membrane pump necessary for urease activity. *Gut* **39** (suppl. 2), A67 (1996).
- McGowan, C. C., Cover, T. L. & Blaser, M. J. *Helicobacter pylori* and gastric acid: biological and therapeutic implications. *Gastroenterology* **110**, 926–938 (1996).
- Hughes, N. J., Chalk, T. L., Clayton, C. L. & Kelly, D. J. Identification of carboxylation enzymes and characterization of a novel four-subunit pyruvate:flavodoxin oxidoreductase from *Helicobacter pylori*. *J. Bacteriol.* **177**, 3953–3959 (1995).
- Mendz, G. L. & Hazell, S. L. Amino acid utilization by *Helicobacter pylori*. *Int. J. Biochem. Cell. Biol.* **27**, 1085–1093 (1995).
- Sonnhammer, E. L. L., Eddy, S. R. & Durbin, R. Pfam: A comprehensive database of protein families based on seed alignments. *Proteins* (in the press).
- Akopyants, N. S., Eaton, K. A. & Berg, D. E. Adaptive mutation and co-colonization during *Helicobacter pylori* infection of gnotobiotic piglets. *Infect. Immun.* **63**, 116–121 (1995).
- Borodovsky, M., Rudd, K. E. & Koonin, E. V. Intrinsic and extrinsic approaches for detecting genes in a bacterial genome. *Nucleic Acids Res.* **22**, 4756–4767 (1994).
- Fuchs, R. MacPattern: protein pattern searching on the Apple Macintosh. *Comput. Appl. Biosci.* **7**, 105–106 (1991).
- Claros, M. G. & von Heijne, G. TopPred II: an improved software for membrane protein structure predictions. *Comput. Appl. Biosci.* **10**, 685–686 (1994).
- Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**, 1–6 (1997).
- Karp, P. D., Riley, M., Paley, S. M., Pellegrini-Toole, A. & Krummenacker, M. EcoCyc: Encyclopedia of *Escherichia coli* genes and metabolism. *Nucleic Acids Res.* **25**, 43–51 (1997).
- Doig, P., Exner, M. M., Hancock, R. E. & Trust, T. J. Isolation and characterization of a conserved porin protein from *Helicobacter pylori*. *J. Bacteriol.* **177**, 5447–5452 (1995).

Acknowledgements. D.E.B., M.B. and W.H. are supported by grants from the NIH; P.K. is supported by a grant from the National Center for Research Resources. We thank N. S. Akopyants for preparing high quality chromosomal DNA from *H. pylori* strain 26695; M. Heaney, J. Scott, A. Saeed and R. Shirley for software and database support; and V. Sapiro, B. Vincent, J. Meehan and D. Mass for computer system support.

Correspondence and requests for materials should be addressed to J.-F.T. (e-mail: ghp@tigr.org). The annotated genome sequence and gene family alignments are available on the World-Wide Web site at <http://www.tigr.org/tdb/mdb/hpdbh/hpdbh.html>. The sequence has been deposited with GenBank under accession number AE000511.



Table 2. List of *H. pylori* genes with putative identifications. Gene numbers correspond to those in Fig. 1. Each identified gene has been assigned a putative role category adapted from ref. 15. Percentages represent per cent identities.

AMINO-ACID BIOSYNTHESIS			CELL ENVELOPE		
<i>General</i>			<i>Membranes, lipoproteins and porins</i>		
HP0695	hydantoin utilization protein A (hyuA)	28.6%	HP0841	pantothenate metabolism flavoprotein (dfp)	31.3%
<i>Aromatic amino-acid family</i>			<i>Pyridoxine</i>		
HP1038	3-dehydroquinate type II (aroC)	99.4%	HP1583	pyridoxal phosphate biosynthetic protein A (pdxA)	34.2%
HP0283	3-dehydroquinate synthase (aroB)	38.1%	HP1582	pyridoxal phosphate biosynthetic protein J (pdxJ)	42.6%
HP0134	3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (dhsl)	54.6%	<i>Riboflavin</i>		
HP0401	3-phosphoshikimate 1-carboxyvinyltransferase (aroA)	53.6%	HP0802	GTP cyclohydrolase II (ribA)	47.2%
HP1279	anthranilate isomerase (trpC)	47.0%	HP0804	GTP cyclohydrolase II/3,4-dihydroxy-2-butanone 4-phosphate synthase (ribA, ribB)	44.0%
HP1282	anthranilate synthase component I (trpE)	47.9%	HP1505	riboflavin biosynthesis protein (ribG)	33.1%
HP1280	anthranilate synthase component II (trpD)	42.5%	HP1087	riboflavin biosynthesis regulatory protein (ribC)	28.9%
HP1281	anthranilate synthase component II (trpD)	40.2%	HP1574	riboflavin synthase alpha subunit (ribC)	32.8%
HP0663	chorismate synthase (aroC)	47.2%	HP0002	riboflavin synthase beta chain (ribE)	52.4%
HP1380	prephenate dehydrogenase (tyrA)	30.2%	<i>Thioresoxin, glutaredoxin and glutathione</i>		
HP1249	shikimate 5-dehydrogenase (aroE)	36.6%	HP1118	gamma-glutamyltranspeptidase (ggT)	53.2%
HP0157	shikimate kinase I (aroK)	36.1%	HP1458	thioredoxin	38.3%
HP1277	tryptophan synthase, alpha subunit (trpA)	46.5%	HP0824	thioredoxin (trxA)	51.5%
HP1278	tryptophan synthase, beta subunit (trpB)	66.1%	HP1164	thioredoxin reductase (trxB)	28.5%
<i>Aspartate family</i>			<i>Thiamine</i>		
HP0649	aspartate ammonia-lyase (aspA)	55.5%	HP0814	thiamin biosynthesis protein (thiF)	34.6%
HP1189	aspartate-semialdehyde dehydrogenase (aspC)	45.7%	HP0843	thiamin phosphate pyrophosphorylase/hydroxyethylthiazole kinase (thiB)	35.7%
HP1229	aspartokinase (lysC)	48.0%	HP0845	thiamin phosphate pyrophosphorylase/hydroxyethylthiazole kinase (thiM)	37.9%
HP0106	cystathionine gamma-synthase (metB)	47.7%	HP0844	thiamine biosynthesis protein (thi)	41.0%
HP0290	diaminopimelate decarboxylase (dap decarboxylase) (lysA)	42.7%	<i>Pyridine nucleotides</i>		
HP0666	diaminopimelate epimerase (dapF)	30.0%	HP0329	NH(3)-dependent NAD+ synthetase (nadE)	37.5%
HP0610	dihydrodipicolinate reductase (dapB)	95.3%	HP1355	nicotinate-nucleotide pyrophosphorylase (nadC)	36.3%
HP1013	dihydrodipicolinate synthetase (dapA)	39.5%	HP1356	quinolinate synthetase A (nadA)	34.2%
HP0822	homoserine dehydrogenase (metL)	37.7%	Surface structures		
HP1050	homoserine kinase (thrB)	27.7%	HP0840	flaA1 protein	60.2%
HP0672	solute-binding signature and mitochondrial signature protein (aspB)	47.3%	HP0325	flagellar basal-body L-ring protein (flgH)	32.7%
HP0212	succinyl-diaminopimelate desuccinylase (dapE)	42.3%	HP0361	flagellar basal-body M-ring protein (flfF)	34.4%
HP0626	tetrahydrodipicolinate N-succinyltransferase (dapD)	36.1%	HP0246	flagellar basal-body P-ring protein (flgI)	37.9%
HP0098	threonine synthase (thrC)	32.9%	HP1557	flagellar basal-body protein (flfE)	37.0%
<i>Glutamate family</i>			HP1559	flagellar basal-body rod protein (flgB) (proximal rod protein)	31.0%
HP0380	glutamate dehydrogenase (gdhA)	58.0%	HP1558	flagellar basal-body rod protein (flgC) (proximal rod protein)	46.0%
HP0512	glutamine synthetase (glnA)	48.6%	HP1092	flagellar basal-body rod protein (flgG)	35.5%
HP1158	pyroline-5-carboxylate reductase (proC)	28.9%	HP1585	flagellar basal-body rod protein (flgF)	47.7%
<i>Pyruvate family</i>			HP1041	flagellar biosynthesis protein (flhA)	43.1%
HP0841	alanine racemase, biosynthetic (alr)	32.4%	HP1035	flagellar biosynthesis protein (flhF)	35.5%
HP1468	branched-chain-amino-acid aminotransferase (ilvE)	63.5%	HP0684	flagellar biosynthesis protein (flhP)	43.4%
HP0330	ketol-acid reductoisomerase (ilvC)	48.1%	HP0770	flagellar biosynthesis protein (flhB)	38.7%
<i>Serine family</i>			HP0685	flagellar biosynthesis protein (flhP)	55.6%
HP0107	cysteine synthetase (cysK)	45.7%	HP1419	flagellar biosynthesis protein (flhI)	52.3%
HP0096	phosphoglycerate dehydrogenase	31.0%	HP0173	flagellar biosynthesis protein (flhR)	26.4%
HP0397	phosphoglycerate dehydrogenase (serA)	32.5%	HP0353	flagellar export protein (flhH)	29.1%
HP0736	phosphoserine aminotransferase (serC)	30.7%	HP1420	flagellar export protein ATP synthase (flhI)	47.6%
HP0652	phosphoserine phosphatase (serB)	36.5%	HP0870	flagellar hook (flgE)	98.9%
HP1210	serine acetyltransferase (cysE)	96.2%	HP0908	flagellar hook (flgE)	30.5%
HP1013	serine hydroxymethyltransferase (glyA)	54.0%	HP1119	flagellar hook-associated protein 1 (HAP1) (flgK)	27.6%
BIOSYNTHESIS OF COFACTORS, PROSTHETIC GROUPS, AND CARRIERS			HP0752	flagellar hook-associated protein 2 (flhD)	28.9%
<i>General</i>			HP0815	flagellar motor rotation protein (motA)	32.9%
HP0220	synthesis of [Fe-S] cluster (nifS)	48.0%	HP0816	flagellar motor rotation protein (motB)	29.7%
<i>Biotin</i>			HP0352	flagellar motor switch protein (flgI)	37.0%
HP0598	8-amino-7-oxononanoate synthase (bioF)	34.9%	HP1031	flagellar motor switch protein (flhM)	34.4%
HP0976	adenosylmethionine-8-amino-7-oxononanoate aminotransferase (bioA)	49.2%	HP0753	flagellar protein (flhS)	32.3%
HP1140	biotin operon repressor/biotin acetyl coenzyme A carboxylase synthetase (birA)	36.9%	HP0327	flagellar protein G (flaG)	23.3%
HP0407	biotin sulfoxide reductase (bioC)	42.7%	HP0797	flagellar sheath adhesin hpaA	98.5%
HP1254	biotin synthetase protein (bioB)	32.1%	HP0584	flagellar switch protein (flhN)	39.7%
HP1406	biotin synthetase (bioB)	36.2%	HP0601	flagellin A (flaA)	99.8%
HP0029	dethiobiotin synthetase (bioD)	36.0%	HP0115	flagellin B (flaB)	99.0%
<i>Folic acid</i>			HP0295	flagellin B homologue (fla)	32.9%
HP1036	7, 8-dihydro-6-hydroxymethylpterin-pyrophosphokinase (folK)	34.6%	HP1575	flhB protein (flhB)	40.5%
HP0587	aminodeoxychorismate lyase (pabC)	32.4%	HP1030	flhY protein (flhY)	29.3%
HP1232	dihydrodipicolinate synthase (folP)	34.5%	HP0907	Hook assembly protein, flagella (flgD)	25.3%
HP1545	7-methylcyclohydrolase (folC)	35.2%	HP1274	paralytic flagellin protein (pflA)	23.9%
HP0928	7-methylcyclohydrolase I (folE)	50.9%	HP0751	paralytic flagellin (flaG)	21.9%
HP0577	methylene-tetrahydrofolate dehydrogenase (folD)	48.4%	HP0410	putative neuraminylactose-binding haemagglutinin homologue (hpaA)	24.2%
HP0293	para-aminobenzoate synthetase (pabB)	35.1%	HP1192	secreted protein involved in flagellar motility	72.5%
<i>Haem and porphyrin</i>			HP1462	secreted protein involved in flagellar motility	96.2%
HP0163	delta-aminolevulinic acid dehydratase (hemB)	50.5%	HP0232	secreted protein involved in flagellar motility	99.2%
HP0376	ferrochelatase (hemH)	33.4%	CELLULAR PROCESSES		
HP0306	glutamate-1-semialdehyde 2,1-aminomutase (hemL)	51.3%	<i>General</i>		
HP0239	glutamyl-tRNA reductase (hemA)	32.7%	HP0018	chemotaxis protein (cheV)	26.8%
HP0665	oxygen-independent coproporphyrinogen III oxidase (hemN)	42.4%	HP0393	chemotaxis protein (cheV)	31.7%
HP1228	oxygen-independent coproporphyrinogen III oxidase (hemN)	37.9%	HP0616	chemotaxis protein (cheV)	27.9%
HP0237	porphobilinogen deaminase (hemC)	45.7%	HP1067	chemotaxis protein (cheV)	99.2%
HP0381	protoporphyrinogen oxidase (hemK)	35.9%	HP0517	GTP-binding protein (era)	95.6%
HP0604	uroporphyrinogen decarboxylase (hemE)	46.3%	HP1490	haemolysin	39.2%
HP1224	uroporphyrinogen III cosynthase (hemD)	27.6%	HP1086	haemolysin (tyl)	40.2%
<i>Menquinone and ubiquinone</i>			HP0599	haemolysin secretion protein precursor (hylB)	45.4%
HP1360	4-hydroxybenzoate octaprenyltransferase (ubiA)	26.6%	HP0392	histidine kinase (cheA)	41.4%
HP0929	gamma-transferrase (ispA)	39.8%	HP0039	methyl-accepting chemotaxis protein (tpaA)	32.8%
HP0240	octaprenyl-diphosphate synthase (ispB)	31.6%	HP1013	methyl-accepting chemotaxis protein (tpaB)	30.7%
<i>Molybdopterins</i>			HP0082	methyl-accepting chemotaxis transducer (tpcC)	28.2%
HP0768	molybdenum cofactor biosynthesis protein A (moaA)	31.4%	HP0391	purine-binding chemotaxis protein (cheW)	34.3%
HP0798	molybdenum cofactor biosynthesis protein C (moaC)	97.9%	<i>Cell division</i>		
HP0172	molybdopterins biosynthesis protein (moaE)	36.3%	HP0331	cell division inhibitor (minD)	50.2%
HP0756	molybdopterins biosynthesis protein (moaB)	32.2%	HP0749	cell division membrane protein (ftsX)	25.7%
HP0799	molybdopterins biosynthesis protein (mog)	50.8%	HP0978	cell division protein (ftsA) protein	31.9%
HP0801	molybdopterins converting factor, subunit 1 (moaD)	31.1%	HP0748	cell division protein (ftsE)	37.6%
HP0800	molybdopterins converting factor, subunit 2 (moaE)	31.1%	HP0286	cell division protein (ftsH)	41.2%
HP0769	molybdopterins-guanine dinucleotide biosynthesis protein A (moaB)	28.3%	HP1089	cell division protein (ftsI)	98.6%
<i>Pantothenate</i>			HP1556	cell division protein (ftsL)	30.6%
HP1058	3-methyl-2-oxobutanate hydroxymethyltransferase (panB)	43.7%	HP1090	cell division protein (ftsK)	39.8%
HP0034	aspartate 1-decarboxylase (panD)	50.0%	HP1560	cell division protein (ftsW) Escherichia coli	32.7%
HP0006	pantoate-beta-alanine ligase (panC)	44.2%	HP0763	cell division protein (ftsY)	46.6%

HP0332	cell division topological specificity factor (minE)	33.8%	HP1270	subunit (NQO10)	-1.0%	HP1101	(devB)	29.2%
HP0979	cell division protein (ftsZ)	43.3%	HP1271	NADH-ubiquinone oxidoreductase, NQO11 subunit (NQO11) (Paracoccus denitrificans)	42.6%	HP1495	glucose-6-phosphate dehydrogenase (g6pD)	36.7%
HP1159	cell filamentation protein (fic)	63.2%	HP1272	NADH-ubiquinone oxidoreductase, NQO12 subunit (NQO12)	43.2%	HP1088	transaldolase (tal)	33.5%
Cell killing			HP1273	NADH-ubiquinone oxidoreductase, NQO13 subunit (NQO13)	40.2%	HP0354	transketolase A (tktA)	46.7%
HP0887	vacuolating cytotoxin	94.7%	HP1266	NADH-ubiquinone oxidoreductase, NQO3 subunit (NQO3)	31.6%	HP0574	transketolase B (tkbA)	39.7%
Chaperones			HP1263	NADH-ubiquinone oxidoreductase, NQO4 subunit (NQO4) (Triticum aestivum)	44.6%	Sugars		
HP0010	chaperone and heat shock protein (groEL)	99.6%	HP1262	NADH-ubiquinone oxidoreductase, NQO5 subunit (NQO5)	-1.0%	HP0360	galactosidase acetyltransferase (lacA)	41.0%
HP0109	chaperone and heat shock protein 70 (dnaK)	63.4%	HP1261	NADH-ubiquinone oxidoreductase, NQO6 subunit (NQO6)	62.2%	HP0779	UDP-glucose 4-epimerase	43.1%
HP0210	chaperone and heat shock protein C62.5 (htpG)	46.5%	HP1260	NADH-ubiquinone oxidoreductase, NQO7 subunit (NQO7)	40.7%	HP0026	aconitase B (aconB)	64.0%
HP0011	co-chaperone (groES)	99.2%	HP1267	NADH-ubiquinone oxidoreductase, NQO8 subunit (NQO8)	42.4%	HP0026	citrate synthase (gltA)	47.8%
HP1332	co-chaperone and heat-shock protein (dnaJ)	42.7%	HP1268	NADH-ubiquinone oxidoreductase, NQO9 subunit (NQO9)	41.2%	HP1325	fumarate (fumC)	63.7%
HP0110	co-chaperone and heat-shock protein (grpE)	33.0%				HP0509	glycolate oxidase subunit (gltD)	98.0%
HP1024	co-chaperone-curved DNA-binding protein A (CbpA)	37.7%				HP0027	isocitrate dehydrogenase (icd)	70.7%
Chromosome-associated protein						FATTY ACID AND PHOSPHOLIPID METABOLISM		
HP1138	plasmid replication-partition related protein	40.4%				General		
Detoxification						HP1376	(3R)-hydroxymyristoyl-(acyl carrier protein) dehydratase (fabZ)	47.4%
HP1563	alkyl hydroperoxide reductase (tsaA)	98.5%	Amino acids and amines			HP1348	1-acyl-glycerol-3-phosphate acyltransferase (plsC) (Escherichia coli)	32.0%
HP0875	catalase	99.4%	HP1398	alanine dehydrogenase (ald)	39.6%	HP0561	3-ketoacyl-acyl carrier protein reductase (fabG)	45.7%
HP0267	chlorohydrilase	42.6%	HP0294	aliphatic amidase (aimE)	75.4%	HP0690	acetyl coenzyme A acetyltransferase (thiolase) (fadA)	52.0%
HP0243	neutrophil activating protein (napA) (bacterioperitin)	95.8%	HP1238	aliphatic amidase (aimE)	37.2%	HP0950	acetyl-CoA carboxylase beta subunit (accD)	49.4%
HP0389	superoxide dismutase (sodB)	98.6%	HP1399	arginase (rocF)	31.8%	HP1045	acetyl-CoA synthetase (accE)	52.3%
HP1452	thiophene and furan oxidizer (tdhF)	37.6%	HP0943	D-amino acid dehydrogenase (dadA)	26.2%	HP0557	acetyl-coenzyme A carboxylase (accA)	50.3%
Protein and peptide secretion			HP0056	delta-1-pyrroline-5-carboxylate dehydrogenase (Synecocystis sp.)	32.2%	HP0559	acyl carrier protein (acpP)	55.3%
HP0355	GTP-binding membrane protein (lepA)	57.3%	HP0723	L-asparaginase II (ansB)	54.1%	HP0962	acyl carrier protein (acpP)	56.3%
HP0074	lipoprotein signal peptidase (lspA)	97.0%	HP0132	L-serine deaminase (sdaA)	45.8%	HP0558	beta ketoacyl-acyl carrier protein synthase II (fabF)	50.0%
HP0786	preprotein translocase subunit (secA)	54.0%	Anaerobic			HP0202	beta-ketoacyl-acyl carrier protein synthase III (fabH)	44.4%
HP1300	preprotein translocase subunit (secY)	41.2%	HP0666	anaerobic glycerol-3-phosphate dehydrogenase, subunit C (glpC)	27.2%	HP0371	biotin carboxyl carrier protein (fabE)	30.8%
HP1255	protein translocation protein, low temperature (secG)	30.6%	HP0589	ferredoxin oxidoreductase, alpha subunit	42.7%	HP0370	biotin carboxylase (accC)	52.1%
HP1550	protein-export membrane protein (secD)	35.9%	HP0680	ferredoxin oxidoreductase, beta subunit	43.2%	HP0871	CDP-diglyceride hydrolase (cdh)	73.9%
HP1549	protein-export membrane protein (secE)	35.1%	HP0591	ferredoxin oxidoreductase, gamma subunit	33.3%	HP0215	CDP-diglyceride synthetase (cdsA)	42.4%
HP0576	signal peptidase I (lepB)	40.3%	HP0193	fumarate reductase, cytochrome b subunit (frcD)	58.8%	HP0416	cyclopropane fatty acid synthase (cfa)	39.7%
HP1152	signal recognition particle protein (fih)	41.4%	HP0192	fumarate reductase, flavoprotein subunit (frcA)	69.4%	HP0700	diacylglycerol kinase (dgaA)	45.8%
HP0795	trigger factor (tig)	27.6%				HP0195	enoyl-(acyl-carrier-protein) reductase (NADH) (fabI)	45.8%
Transformation						HP0201	fatty acid/phospholipid synthesis protein (plsX)	37.8%
HP0520	cag pathogenicity island protein (cag1)	96.5%	HP0191	fumarate reductase, iron-sulfur subunit (frcB)	70.8%	HP0808	Holo-acp synthase (acpS)	29.1%
HP0530	cag pathogenicity island protein (cag10)	98.4%	HP1110	pyruvate ferredoxin oxidoreductase, alpha subunit	41.0%	HP0090	malonyl coenzyme A-acyl carrier protein transacylase (fabD)	35.4%
HP0531	cag pathogenicity island protein (cag11)	97.2%	HP1111	pyruvate ferredoxin oxidoreductase, beta subunit	43.7%	HP1016	phosphatidylglycerophosphate synthase (pgsA)	35.4%
HP0532	cag pathogenicity island protein (cag12)	98.9%	HP1109	pyruvate ferredoxin oxidoreductase, delta subunit	47.0%	HP1357	phosphatidylserine decarboxylase proenzyme (psd)	33.2%
HP0534	cag pathogenicity island protein (cag13)	98.0%	HP1108	pyruvate ferredoxin oxidoreductase, gamma subunit	37.2%	HP1071	phosphatidylserine synthase (psaA)	99.6%
HP0535	cag pathogenicity island protein (cag14)	97.6%	ATP-protonmotive force interconversion			HP0499	phospholipase A1 precursor (DR-phospholipase A)	33.8%
HP0536	cag pathogenicity island protein (cag15)	96.4%	HP0828	ATP synthase FO, subunit a (atpB)	37.7%	PURINES, PYRIMIDINES, NUCLEOSIDES AND NUCLEOTIDES		
HP0537	cag pathogenicity island protein (cag16)	98.9%	HP1136	ATP synthase FO, subunit b (atpF)	28.3%	General		
HP0538	cag pathogenicity island protein (cag17)	95.3%	HP1137	ATP synthase FO, subunit b0 (atpF0)	32.5%	HP0757	beta-alanine synthetase homologue	40.0%
HP0539	cag pathogenicity island protein (cag18)	98.7%	HP1212	ATP synthase FO, subunit c (atpE)	41.2%	2'-Deoxyribonucleotide metabolism		
HP0540	cag pathogenicity island protein (cag19)	99.5%	HP1134	ATP synthase F1, subunit alpha (atpA)	62.7%	HP0372	deoxycytidine triphosphate deaminase (dcd)	28.2%
HP0521	cag pathogenicity island protein (cag20)	92.5%	HP1132	ATP synthase F1, subunit beta (atpD)	85.6%	HP0865	deoxycytidine 5'-triphosphate nucleotidohydrolase (dnt)	41.4%
HP0541	cag pathogenicity island protein (cag21)	97.8%	HP1135	ATP synthase F1, subunit delta (atpH)	24.6%	HP0364	ribonucleoside diphosphate reductase, beta subunit (nrdB)	39.0%
HP0542	cag pathogenicity island protein (cag22)	95.5%	HP1131	ATP synthase F1, subunit epsilon (atpC)	32.7%	HP0680	ribonucleoside-diphosphate reductase 1 alpha subunit (nrdA)	28.4%
HP0544	cag pathogenicity island protein (cag23)	99.0%	HP1133	ATP synthase F1, subunit gamma (atpG)	37.8%	HP0825	thioredoxin reductase (trxB)	45.9%
HP0545	cag pathogenicity island protein (cag24)	98.5%	Electron transport			Purine ribonucleotide biosynthesis		
HP0546	cag pathogenicity island protein (cag25)	95.7%	HP0146	cb3-type cytochrome c oxidase subunit Q (CooQ)	44.2%	HP0321	5'-guanylate kinase (gmk)	44.8%
HP0547	cag pathogenicity island protein (cag26)	92.9%	HP0265	cytochrome c biogenesis protein (ccdA)	35.4%	HP0618	adenylate kinase (ack)	33.3%
HP0522	cag pathogenicity island protein (cag3)	98.1%	HP0378	cytochrome c biogenesis protein (cyf5)	37.5%	HP1112	adenylosuccinate lyase (purB)	49.5%
HP0523	cag pathogenicity island protein (cag4)	95.7%	HP0147	cytochrome c oxidase, dihelme subunit, membrane-bound (flxP)	33.0%	HP0255	adenylosuccinate synthase (purA)	44.6%
HP0524	cag pathogenicity island protein (cag5)	99.1%	HP0144	cytochrome c oxidase, heme b and copper-binding subunit, membrane-bound (flxB)	43.9%	HP1434	formyltetrahydrofolate hydrolase (purU)	49.1%
HP0525	cag pathogenicity island protein (cag6)	97.5%	HP0145	cytochrome c oxidase, monoheme subunit, membrane-bound (flxO)	45.7%	HP1218	glycinamide ribonucleotide synthetase (purD)	31.8%
HP0527	cag pathogenicity island protein (cag7)	94.8%	HP1461	cytochrome c551 peroxidase	48.5%	HP0854	GMP reductase (guaC)	31.8%
HP0528	cag pathogenicity island protein (cag8)	99.0%	HP1227	cytochrome c553	38.4%	HP0409	GMP synthase (guaA)	56.1%
HP0529	cag pathogenicity island protein (cag9)	98.9%	HP0277	ferredoxin	52.5%	HP0829	inosine-5'-monophosphate dehydrogenase (guaB)	58.5%
HP1378	competence lipoprotein (comL)	25.5%	HP0588	ferredoxin-like protein	42.6%	HP0198	nucleoside diphosphate kinase (ndk)	67.7%
HP1361	competence locus E (comE3)	26.7%	HP1508	ferredoxin-like protein	29.4%	HP0742	phosphoribosylpyrophosphate synthetase (prsA)	56.5%
HP1006	conjugal transfer protein (traG)	27.3%	HP1161	flavodoxin (fldA)	47.0%	HP1530	purine nucleoside phosphorylase (punB)	20.7%
HP1421	conjugative transfer region protein (trbB)	30.7%	HP0642	NAD(P)H-flavin oxidoreductase	46.1%	Pyrimidine ribonucleotide biosynthesis		
HP0533	DNA processing chain A (dprA)	32.9%	HP0954	oxygen-insensitive NAD(P)H nitroreductase	32.7%	HP1084	aspartate transcarbamoylase (pyrB)	38.7%
HP0042	trb1 protein	31.4%	HP0634	quinone-reactive Ni/Fe hydrogenase (hydD)	54.7%	HP0919	carbamoyl-phosphate synthase (glutamine-hydrolysing) (pyrAb)	48.6%
HP0525	VirB11 homologue	100.0%	HP0633	quinone-reactive Ni/Fe hydrogenase, cytochrome b subunit (hydC)	51.4%	HP1237	carbamoyl-phosphate synthetase (pyrAa)	39.7%
HP0441	VirB4 homologue	23.5%	HP0632	quinone-reactive Ni/Fe hydrogenase, large subunit (hydE)	68.5%	HP0349	GTP synthetase (pyrG)	50.7%
HP0017	virB4 homologue (virB4)	25.2%	HP0631	quinone-reactive Ni/Fe hydrogenase, small subunit (hydA)	68.9%	HP0266	dihydroorotase (pyrC)	-1.0%
HP0459	virB4 homologue (virB4)	25.3%	HP1539	ubiquinol cytochrome c oxidoreductase, cytochrome b subunit (fbcH)	39.3%	HP0581	dihydroorotase (pyrC)	31.5%
CENTRAL INTERMEDIARY METABOLISM			HP1538	ubiquinol cytochrome c oxidoreductase, cytochrome c1 subunit (fbcH)	28.8%	HP1011	dihydroorotate dehydrogenase (pyrD)	41.6%
General			HP1540	ubiquinol cytochrome c oxidoreductase, Rieske 2Fe-2S subunit (fbcF)	39.2%	HP1257	orotate phosphoribosyltransferase (pyrE)	35.5%
HP1014	7-alpha-hydroxysteroid dehydrogenase (hdhA)	33.2%	Entner-Doudoroff			HP0005	orotidine 5'-phosphate decarboxylase (pyrF)	39.0%
HP1186	carbonic anhydrase	37.0%	HP1099	2-keto-3-deoxy-6-phosphogluconate aldolase (eda)	50.3%	HP1474	thymidylate kinase (tmk)	33.9%
HP0004	carbonic anhydrase (icfA)	33.3%	HP1100	6-phosphogluconate dehydratase	50.7%	HP0777	uridine 5'-monophosphate (UMP) kinase (pyrH)	50.4%
HP0869	hydrogenase expression/formation protein (hypA)	28.1%				Salvage of nucleosides and nucleotides		
HP0900	hydrogenase expression/formation protein (hypB)	41.4%				HP1014	2'-O-cyclic-nucleotide 2'-phosphodiesterase (cptB)	31.8%
HP0899	hydrogenase expression/formation protein (hypC)	38.5%				HP0572	adenine phosphoribosyltransferase (apt)	50.3%
HP0898	hydrogenase expression/formation protein (hypD)	47.8%				HP1179	phosphotransferase (deoB)	55.9%
HP0047	hydrogenase expression/formation protein (hypE)	39.7%				HP1178	purine-nucleoside phosphorylase (deoD)	55.5%
HP0197	S-adenosylmethionine synthetase 2 (metX)	62.1%				HP0735	xanthine guanine phosphoribosyl transferase (gpt)	27.1%
Amino sugars						Sugar-nucleotide biosynthesis and conversions		
HP1532	glucosamine fructose-6-phosphate aminotransferase (isomerizing) (glmS)	41.7%				HP0043	mannose-6-phosphate isomerase (pmi) or (algA)	42.8%
Phosphorus compounds						HP0045	modulation protein (nolK)	44.3%
HP0620	inorganic pyrophosphatase (ppa)	50.0%				HP0646	UDP-glucose pyrophosphorylase (galU)	65.6%
HP0696	N-methylhydantoinase	26.9%				HP0683	UDP-N-acetylglucosamine pyrophosphorylase (glmU)	40.0%
HP1010	polyphosphate kinase (ppk)	38.5%				REGULATORY FUNCTIONS		
Polyamine biosynthesis						General		
HP0422	arginine decarboxylase (speA)	33.3%				HP1032	alternative transcription initiation factor, sigma-F (flaA)	34.6%
HP0020	carboxynorspermidine decarboxylase (nspC)	45.6%				HP1168	carbon starvation protein (cstA)	59.8%
HP0832	spermidine synthase (speE)	26.5%				HP1442	carbon storage regulator (cstA)	43.3%
Other						HP1027	feric uptake regulation protein (fur)	39.9%
HP0070	urease accessory protein (ureF)	97.1%				HP0278	guanosine pentaphosphate phosphohydrolase (gppA)	26.4%
HP0069	urease accessory protein (ureG)	94.5%				HP0400	penicillin tolerance protein (lytB)	30.6%
HP0068	urease accessory protein (ureH)	95.0%						
HP0067	urease accessory protein (ureI)	96.2%						
HP0071	urease accessory protein (ureJ)	98.5%						
HP0073	urease alpha subunit (ureA)	100.0%						
HP0072	urease beta subunit (urea amidohydrolase) (ureB)	100.0%						
HP0075	urease protein (ureC)	98.0%						
ENERGY METABOLISM								
Aerobic								
HP1222	D-lactate dehydrogenase (ldd)	27.0%						
HP0961	glycerol-3-phosphate dehydrogenase (NAD(P)H)	36.8%						
HP0037	NADH-ubiquinone oxidoreductase subunit	19.4%						
HP1269	NADH-ubiquinone oxidoreductase, NQO10							

HP0775	penta-phosphate guanosine-3O-pyrophosphohydrolase (spoT)	36.7%	HP1471	type IIS restriction enzyme R protein (BcglRI)	28.2%	HP0399	ribosomal protein S1 (rps1)	30.5%
HP0224	peptide methionine sulphoxide reductase (msrA)	66.8%	HP1366	type IIS restriction enzyme R protein (MBOIR)	37.1%	HP1320	ribosomal protein S10 (rps10)	58.2%
HP1025	putative heat shock protein (hspR)	46.2%	HP1208	ulcer associated adenine specific DNA methyltransferase	93.4%	HP1235	ribosomal protein S11 (rps11)	56.2%
HP1572	regulatory protein DnrI	31.9%	HP1209	ulcer-associated gene restriction endonuclease (iceA)	95.5%	HP1197	ribosomal protein S12 (rps12)	79.0%
HP0703	response regulator	44.2%	HP1347	uracil-DNA glycosylase (ung)	43.1%	HP1296	ribosomal protein S13 (rps13)	55.8%
HP1021	response regulator	28.7%	TRANSCRIPTION		HP1306	ribosomal protein S14 (rps14)	68.3%	
HP1043	response regulator	26.8%	Degradation of RNA		HP1040	ribosomal protein S15 (rps15)	57.8%	
HP1365	response regulator	32.4%	HP1213	polynucleotide phosphorylase (pnp)	38.9%	HP1151	ribosomal protein S16 (rps16)	46.8%
HP0166	response regulator (ompR)	51.0%	HP1293	DNA-dependent RNA polymerase		HP1310	ribosomal protein S17 (rps17)	55.4%
HP0714	RNA polymerase sigma-54 factor (rpoN)	37.1%	HP1198	DNA-directed RNA polymerase, alpha subunit (rpoA)	35.3%	HP1244	ribosomal protein S18 (rps18)	55.2%
HP0088	RNA polymerase sigma-70 factor (rpoD)	43.5%	HP1198	DNA-directed RNA polymerase, beta subunit (rpoB)	47.8%	HP1315	ribosomal protein S19 (rps19)	61.1%
HP0792	sigma-54 interacting protein	97.7%	Transcription factors		HP1554	ribosomal protein S2 (rps2)	49.6%	
HP0164	signal-transducing protein, histidine kinase	27.1%	HP0866	transcription elongation factor GreA (greA)	50.3%	HP0076	ribosomal protein S20 (rps20)	41.4%
HP1364	signal-transducing protein, histidine kinase	24.9%	HP1514	transcription termination factor NusA (nusA)	39.1%	HP0662	ribosomal protein S21 (rps21)	42.4%
HP0244	signal-transducing protein, histidine kinase (atoS)	30.0%	HP0001	transcription termination factor NusB (nusB)	30.2%	HP1313	ribosomal protein S3 (rps3)	56.7%
HP0048	transcriptional regulator (hvpF)	34.5%	HP1203	transcription termination factor NusG (nusG)	41.0%	HP1294	ribosomal protein S4 (rps4)	51.2%
HP1287	transcriptional regulator (tenA)	34.7%	HP0550	transcription termination factor Rho (rho)	56.6%	HP1302	ribosomal protein S5 (rps5)	65.5%
HP0727	transcriptional regulator, putative	33.3%	RNA processing		HP1246	ribosomal protein S6 (rps6)	32.1%	
REPLICATION			TRANSLATION		HP1196	ribosomal protein S7 (rps7)	62.2%	
Degradation of DNA			General		HP1305	ribosomal protein S8 (rps8)	45.0%	
HP0275	ATP-dependent nuclease (addB)	27.2%	HP0944	translation initiation inhibitor, putative	45.6%	HP0083	ribosomal protein S9 (rps9)	50.4%
HP0259	exonuclease VII, large subunit (xseA)	37.6%	Aminoacyl tRNA synthetases		HP1047	ribosome-binding factor A (rbfA)	26.3%	
DNA replication, restriction, modification, recombination and repair			Alanyl-tRNA synthetase (alaS)		tRNA modification			
HP0142	A/G-specific adenine glycosylase (mutY)	38.2%	HP0241	alanyl-tRNA synthetase (alaS)	44.9%	HP1141	methionyl-tRNA formyltransferase (fmt)	37.5%
HP0050	adenine specific DNA methyltransferase (dpmA)	37.4%	HP0319	arginyl-tRNA synthetase (argS)	35.8%	HP1497	peptidyl-tRNA hydrolase (pth)	46.6%
HP0910	adenine specific DNA methyltransferase (HINDIII)	33.4%	HP0617	aspartyl-tRNA synthetase (aspS)	50.1%	HP0361	pseudouridylyl synthase I (hisT)	32.2%
HP1352	adenine specific DNA methyltransferase (HINFIM)	62.5%	HP0886	cysteinyl-tRNA synthetase (cysS)	97.3%	HP1448	ribonuclease P, protein component (mpaA)	29.3%
HP0263	adenine specific DNA methyltransferase (hpaim)	33.9%	HP0476	glutamyl-tRNA synthetase (glxS)	43.1%	HP1062	S-adenosylmethionine:tRNA	
HP0481	adenine specific DNA methyltransferase (MFOKI)	29.3%	HP0643	glutamyl-tRNA synthetase (glxS)	39.8%	HP1513	ribosyltransferase-isomerase (queA)	39.3%
HP0260	adenine specific DNA methyltransferase (mod)	33.9%	HP0960	glycyl-tRNA synthetase, alpha subunit (glyG)	60.1%	HP1148	selenocystein synthase (selA)	36.2%
HP0593	adenine specific DNA methyltransferase (mod)	38.5%	HP0972	glycyl-tRNA synthetase, beta subunit (glyS)	60.1%	HP1418	tRNA (guanine-N1)-methyltransferase (trmD)	39.1%
HP1522	adenine specific DNA methyltransferase (mod)	42.2%	HP1190	histidyl-tRNA synthetase (hisS)	32.4%	HP1415	tRNA delta(2)-isopentenylpyrophosphate transferase (miaA)	30.7%
HP0478	adenine specific DNA methyltransferase (VSPIM)	42.1%	HP1422	isoleucyl-tRNA synthetase (ileS)	49.7%	HP0281	tRNA-guanine transglycosylase (tgt)	45.6%
HP0054	adenine/cytosine DNA methyltransferase	32.1%	HP1547	leucyl-tRNA synthetase (leuS)	45.9%	Translation factors		
HP0790	anti-codon nuclease masking agent (prfB)	42.9%	HP0182	lysyl-tRNA synthetase (lysS)	58.6%	HP0247	ATP-dependent RNA helicase, DEAD-box family (deaD)	37.7%
HP1529	chromosomal replication initiator protein (dnaA)	34.9%	HP0417	methionyl-tRNA synthetase (metS)	42.4%	HP0077	peptide chain release factor RF-1 (prfA)	52.6%
HP1121	cytosine specific DNA methyltransferase (BSP6IM)	37.0%	HP0403	phenylalanyl-tRNA synthetase, alpha subunit (pheS)	48.7%	HP0171	peptide chain release factor RF-2 (prfB)	49.6%
HP0051	cytosine specific DNA methyltransferase (DDEM)	39.0%	HP0402	phenylalanyl-tRNA synthetase, beta subunit (pheT)	30.0%	HP1256	ribosome releasing factor (lrr)	43.7%
HP0483	cytosine specific DNA methyltransferase (HPHIMC)	38.7%	HP0238	prolyl-tRNA synthetase (proS)	39.8%	HP1195	translation elongation factor EF-G (fusA)	67.5%
HP0701	DNA gyrase, sub A (gyrA)	97.4%	HP1480	seryl-tRNA synthetase (serS)	48.3%	HP0177	translation elongation factor EF-P (efp)	45.1%
HP0601	DNA gyrase, sub B (gyrB)	46.0%	HP1253	threonyl-tRNA synthetase (thrS)	42.1%	HP1555	translation elongation factor EF-Ts (tsf)	43.1%
HP1478	DNA helicase I (uvrD)	35.3%	HP0774	tyrosyl-tRNA synthetase (tyrS)	54.7%	HP1205	translation elongation factor EF-Tu (tufB)	89.5%
HP0548	DNA helicase, putative	38.8%	HP1153	valyl-tRNA synthetase (valS)	43.7%	HP1298	translation initiation factor EF-1 (infA)	65.3%
HP0615	DNA ligase (lig)	40.1%	Degradation of proteins, peptides and glycopeptides		HP1048	translation initiation factor IF-2 (infB)	45.4%	
HP0621	DNA mismatch repair protein (MutS)	32.6%	HP0570	aminopeptidase a1 (pepA)	38.5%	HP0124	translation initiation factor IF-3 (infC)	43.4%
HP1470	DNA polymerase I (polA)	40.0%	HP0033	ATP-dependent Clp protease (clpA)	40.3%	TRANSPORT AND BINDING PROTEINS		
HP1460	DNA polymerase III alpha-subunit (dnaE)	42.0%	HP0794	ATP-dependent clp protease proteolytic component (clpP)	64.6%	General		
HP0500	DNA polymerase III beta-subunit (dnaH)	26.0%	HP1379	ATP-dependent protease (lon)	43.9%	HP0179	ABC transporter, ATP-binding protein	66.7%
HP1231	DNA polymerase III delta prime subunit (dnaE)	48.6%	HP0223	ATP-dependent protease (sms)	41.0%	HP0613	ABC transporter, ATP-binding protein	31.1%
HP1387	DNA polymerase III epsilon subunit (dnaQ)	35.1%	HP1374	ATP-dependent protease ATPase subunit (clpX)	56.3%	HP0715	ABC transporter, ATP-binding protein	52.3%
HP0717	DNA polymerase III gamma and tau subunits (dnaX)	39.0%	HP0264	ATP-dependent protease binding subunit (clpB)	97.7%	HP1576	ABC transporter, ATP-binding protein (abc)	48.2%
HP0012	DNA primase (dnaG)	36.6%	HP0169	collagenase (prtC)	40.1%	HP1465	ABC transporter, ATP-binding protein (HI087)	37.8%
HP1523	DNA recombinase (recG)	32.7%	HP0516	heat-shock protein (hspA) ORF1	98.4%	HP1220	ABC transporter, ATP-binding protein (yhocG)	31.5%
HP1393	DNA repair protein (recN)	28.3%	HP0051	heat-shock protein (hspA)	57.1%	HP0653	ABC transporter, ATP-binding protein (yheS)	36.3%
HP0716	DNA topoisomerase I (topA)	45.1%	HP0470	oligodeoxyphosphatase F (pepF)	97.9%	HP0657	ABC transporter, permease protein (yaeE)	43.1%
HP0440	DNA topoisomerase I (topA)	31.7%	HP0657	processing protease (ymxG)	24.2%	HP0607	acetylcholine resistance protein (acrB)	29.7%
HP0602	endonuclease III	36.6%	HP1485	proline dipeptidase (pepG)	35.2%	HP1432	histidine and glutamine-rich protein	50.0%
HP0885	endonuclease III (nth)	40.1%	HP1350	protease	40.6%	HP1427	histidine-rich, metal binding polypeptide (hpn)	100.0%
HP0705	exonuclease ABC subunit A (uvrA)	53.4%	HP1012	protease (pqeE)	29.6%	HP1206	multidrug-resistance protein (hetA)	26.2%
HP1114	exonuclease ABC subunit B (uvrB)	53.1%	HP1435	protease IV (PspA)	41.7%	HP1082	multidrug-resistance protein (msbA)	32.4%
HP0821	exonuclease ABC subunit C (uvrC)	31.5%	HP0404	protein kinase C inhibitor (SP-P16436)	40.2%	HP0600	multidrug-resistance protein (spaB)	29.7%
HP1526	exodeoxyribonuclease (lexA)	58.9%	HP1019	serine protease (htrA)	52.9%	HP1181	multidrug-efflux transporter	29.1%
HP0213	glucose inhibited division protein (gidA)	48.5%	HP1584	signalycoprotease (gcp)	35.7%	HP0497	sodium- and chloride-dependent transporter	31.7%
HP1063	glucose-inhibited division protein (gidB)	32.9%	HP0382	zinc-metalloprotease (YJR117W)	36.2%	HP0498	sodium- and chloride-dependent transporter	30.8%
HP1563	Holliday junction DNA helicase (ruvA)	39.0%	Nucleoproteins		HP0214	sodium-dependent transporter (huNadC-1)	36.6%	
HP1059	Holliday junction DNA helicase (ruvB)	54.6%	HP0835	histone-like DNA-binding protein HU (hup)	44.6%	Amino acids, peptides and amines		
HP0877	Holliday junction endodeoxyribonuclease (ruvC)	34.7%	Protein modification		HP0940	amino acid ABC transporter, periplasmic binding protein (yckK)	41.5%	
HP0675	integrase/recombinase (xerC)	31.8%	HP0383	L-isocaparytyl-protein carboxyl methyltransferase (pcm)	43.0%	HP0939	amino acid ABC transporter, permease protein (yckL)	46.9%
HP0995	integrase/recombinase (xerD)	27.8%	HP1299	methionine amino peptidase (map)	43.0%	HP1017	amino acid permease (rocE)	41.7%
HP0323	membrane bound endonuclease (nuc)	31.1%	HP1441	peptidyl-prolyl cis-trans isomerase B, cyclosporin-type rotamase (ppl)	58.1%	HP0942	D-alanine glycine permease (dagA)	44.5%
HP0678	methylated-DNA-protein-cysteine methyltransferase (dat1)	41.0%	HP1123	peptidyl-prolyl cis-trans isomerase, FKBP-type rotamase (ylyC)	40.4%	HP0301	dipeptide ABC transporter, ATP-binding protein (dppD)	59.5%
HP0387	primosomal protein replication factor (priA)	36.3%	HP0793	polypeptide deformylase (dof)	41.8%	HP0302	dipeptide ABC transporter, ATP-binding protein (dppF)	54.8%
HP0153	recombinase (recA)	99.1%	Ribosomal proteins: synthesis and modification		HP0298	dipeptide ABC transporter, periplasmic dipeptide-binding protein (dppA)	39.8%	
HP0925	recombinational DNA repair protein (recR)	36.5%	HP1201	ribosomal protein L1 (rpl1)	52.0%	HP0299	dipeptide ABC transporter, permease protein (dppB)	49.3%
HP0911	rep helicase, single-stranded DNA-dependent ATPase (rep)	33.8%	HP1200	ribosomal protein L10 (rpl10)	30.4%	HP0300	dipeptide ABC transporter, permease protein (dppC)	52.5%
HP1362	replicative DNA helicase (dnaB)	39.4%	HP1202	ribosomal protein L11 (rpl11)	63.8%	HP1506	glutamate permease (gltS)	56.9%
HP1383	restriction modification system S subunit	38.1%	HP1088	ribosomal protein L11 methyltransferase (prmA)	38.4%	HP1171	glutamine ABC transporter, ATP-binding protein (glnQ)	51.9%
HP0661	ribonuclease H (rnhA)	58.4%	HP0084	ribosomal protein L13 (rpl13)	50.0%	HP1172	glutamine ABC transporter, periplasmic glutamine-binding protein (glnH)	32.2%
HP1323	ribonuclease HII (rnhB)	36.3%	HP1309	ribosomal protein L14 (rpl14)	65.9%	HP1169	glutamine ABC transporter, permease protein (glnP)	27.6%
HP1245	single-strand DNA-binding protein (ssb)	32.6%	HP1301	ribosomal protein L15 (rpl15)	42.5%	HP1170	glutamine ABC transporter, permease protein (glnP)	30.9%
HP0348	single-stranded-DNA-specific exonuclease (recI)	33.6%	HP1312	ribosomal protein L16 (rpl16)	62.4%	HP0250	oligopeptide ABC transporter, ATP-binding protein (oppD)	39.1%
HP1009	site-specific recombinase	21.3%	HP1232	ribosomal protein L17 (rpl17)	48.3%	HP1252	oligopeptide ABC transporter, periplasmic oligopeptide-binding protein (oppA)	28.7%
HP1541	transcription-repair coupling factor (trcF)	37.7%	HP1303	ribosomal protein L18 (rpl18)	45.5%	HP1251	oligopeptide ABC transporter, permease protein (oppB)	59.6%
HP0462	type I restriction enzyme S protein (hsdS)	37.0%	HP1147	ribosomal protein L19 (rpl19)	50.9%	HP0251	oligopeptide ABC transporter, permease protein (oppC)	31.4%
HP0463	type I restriction enzyme M protein (hsdM)	29.4%	HP0126	ribosomal protein L20 (rpl20)	54.8%	HP0819	osmoprotection protein (proV)	38.3%
HP0464	type I restriction enzyme R protein (hsdR)	31.7%	HP0296	ribosomal protein L21 (rpl21)	46.1%	HP0818	osmoprotection protein (proWX)	30.4%
HP0846	type I restriction enzyme M protein (hsdM)	48.0%	HP1314	ribosomal protein L22 (rpl22)	44.9%	HP0055	proline permease (putP)	51.4%
HP0848	type I restriction enzyme S protein (hsdS)	37.0%	HP1317	ribosomal protein L23 (rpl23)	31.7%	HP0936	proline/betaine transporter (proP)	29.1%
HP0850	type I restriction enzyme M protein (hsdM)	54.4%	HP1308	ribosomal protein L24 (rpl24)	52.2%	HP0133	serine transporter (sdcA)	44.6%
HP1402	type I restriction enzyme R protein (hsdR)	26.6%	HP0297	ribosomal protein L27 (rpl27)	64.7%	Anions		
HP1403	type I restriction enzyme M protein (hsdM)	37.1%	HP0491	ribosomal protein L28 (rpl28)	41.7%	HP0476	molybdenum ABC transporter, ATP-binding protein (modD)	38.4%
HP1404	type I restriction enzyme S protein (hsdS)	36.0%	HP1311	ribosomal protein L29 (rpl29)	45.6%	HP0473	molybdenum ABC transporter, periplasmic molybdate-binding protein (modA)	95.9%
HP0092	type II restriction enzyme M protein (hsdM)	55.3%	HP1319	ribosomal protein L3 (rpl3)	41.8%	HP0474	molybdenum ABC transporter, permease protein (modB)	28.7%
HP0091	type II restriction enzyme R protein (hsdR)	50.7%	HP0551	ribosomal protein L31 (rpl31)	49.3%	HP0313	nitrite extrusion protein (narK)	23.6%
HP1369	type III restriction enzyme M protein (mod)	45.6%	HP0200	ribosomal protein L32 (rpl32)	41.7%	HP1491	phosphate permease	34.8%
HP1370	type III restriction enzyme M protein (mod)	37.0%	HP1204	ribosomal protein L33 (rpl33)	55.1%	Carbohydrates, organic alcohols and acids		
HP1371	type III restriction enzyme R protein	26.2%	HP1447	ribosomal protein L34 (rpl34)	70.5%	HP1043	2-oxoglutarate/malate translocator (SODT1)	37.0%
HP0592	type III restriction enzyme R protein (res)	30.6%	HP0125	ribosomal protein L35 (rpl35)	50.8%	HP1091	alpha-ketoglutarate permease (kgpP)	45.9%
HP1521	type III restriction enzyme R protein (res)	33.1%	HP1297	ribosomal protein L36 (rpl36)	81.6%	HP0724	anaerobic C4-dicarboxylate transport protein (dcaA)	53.8%
HP1472	type IIS restriction enzyme M protein (mod)	32.4%	HP1318	ribosomal protein L4 (rpl4)	40.6%	HP1174	glucose/galactose transporter (gluP)	53.8%
HP1367	type IIS restriction enzyme M1 protein (mod) (Moraxella bovis)	59.3%	HP1307	ribosomal protein L5 (rpl5)	53.1%	HP0141	L-lactate permease (lcpP)	55.5%
HP1368	type IIS restriction enzyme M2 protein (mod)	33.0%	HP1304	ribosomal protein L6 (rpl6)	44.4%	HP0140	L-lactate permease (lcpP)	58.7%
HP1517	type IIS restriction enzyme R and M protein (ECO57IR)	26.7%	HP1199	ribosomal protein L7/L12 (rpl7/l12)	65.0%			
			HP0514	ribosomal protein L9 (rpl9)	39.6%			

Cations			HP0258	conserved hypothetical integral membrane protein	32.7%	HP0728	conserved hypothetical protein	29.3%
HP0791	cadmium-transporting ATPase, P-type (cdaA)	97.5%	HP0284	conserved hypothetical integral membrane protein	29.2%	HP0734	conserved hypothetical protein	31.0%
HP0969	cation efflux system protein (czcA)	37.3%	HP0362	conserved hypothetical integral membrane protein	28.8%	HP0745	conserved hypothetical protein	33.7%
HP1328	cation efflux system protein (czcA)	28.9%	HP0415	conserved hypothetical integral membrane protein	44.4%	HP0747	conserved hypothetical protein	32.4%
HP1329	cation efflux system protein (czcA)	31.3%	HP0467	conserved hypothetical integral membrane protein	100.0%	HP0760	conserved hypothetical protein	36.1%
HP1503	cation-transporting ATPase, P-type (copA)	30.3%	HP0571	conserved hypothetical integral membrane protein	29.5%	HP0810	conserved hypothetical protein	31.0%
HP1073	copper ion binding protein (copP)	92.4%	HP0644	conserved hypothetical integral membrane protein	30.3%	HP0813	conserved hypothetical protein	32.5%
HP1072	copper-transporting ATPase, P-type (copA)	93.9%	HP0677	conserved hypothetical integral membrane protein	28.5%	HP0823	conserved hypothetical protein	27.8%
HP0471	glutathione-regulated potassium-efflux system protein (ketB)	99.3%	HP0693	conserved hypothetical integral membrane protein	46.7%	HP0860	conserved hypothetical protein	52.1%
HP0687	iron(II) transport protein (fecB)	33.6%	HP0718	conserved hypothetical integral membrane protein	33.5%	HP0890	conserved hypothetical protein	32.2%
HP1561	iron(III) ABC transporter, periplasmic iron-binding protein (cseE)	27.5%	HP0737	conserved hypothetical integral membrane protein	33.3%	HP0891	conserved hypothetical protein	33.9%
HP1562	iron(III) ABC transporter, periplasmic iron-binding protein (cseE)	28.2%	HP0758	conserved hypothetical integral membrane protein	47.6%	HP0892	conserved hypothetical protein	39.1%
HP0888	iron(III) dicitrate ABC transporter, ATP-binding protein (fecC)	34.4%	HP0759	conserved hypothetical integral membrane protein	31.1%	HP0894	conserved hypothetical protein	39.8%
HP0889	iron(III) dicitrate ABC transporter, permease protein (fecD)	38.3%	HP0787	conserved hypothetical integral membrane protein	25.2%	HP0926	conserved hypothetical protein	30.7%
HP0686	iron(III) dicitrate transport protein (fecA)	29.7%	HP0851	conserved hypothetical integral membrane protein	37.3%	HP0934	conserved hypothetical protein	33.6%
HP0807	iron(III) dicitrate transport protein (fecA)	28.5%	HP0920	conserved hypothetical integral membrane protein	36.3%	HP0956	conserved hypothetical protein	36.2%
HP1400	iron(III) dicitrate transport protein (fecA)	26.3%	HP0946	conserved hypothetical integral membrane protein	35.9%	HP0959	conserved hypothetical protein	31.1%
HP1344	magnesium and cobalt transport protein (corA)	26.3%	HP0962	conserved hypothetical integral membrane protein	38.5%	HP0966	conserved hypothetical protein	29.1%
HP1183	Na ⁺ /H ⁺ antiporter (napA)	26.6%	HP0983	conserved hypothetical integral membrane protein	32.8%	HP0975	conserved hypothetical protein	25.0%
HP1552	Na ⁺ /H ⁺ antiporter (nhaA)	49.2%	HP1044	conserved hypothetical integral membrane protein	30.6%	HP1020	conserved hypothetical protein	31.5%
HP1077	nickel transport protein (nixA)	98.7%	HP1061	conserved hypothetical integral membrane protein	35.0%	HP1037	conserved hypothetical protein	96.9%
HP0490	putative potassium channel protein, putative	25.7%	HP1080	conserved hypothetical integral membrane protein	44.0%	HP1046	conserved hypothetical protein	32.6%
Nucleosides, purines and pyrimidines			HP1162	conserved hypothetical integral membrane protein	27.6%	HP1049	conserved hypothetical protein	39.7%
HP1290	nicotinamide mononucleotide transporter (pnuC)	28.0%	HP1175	conserved hypothetical integral membrane protein	40.6%	HP1066	conserved hypothetical protein	41.3%
HP1180	pyrimidine nucleoside transport protein (nupC)	32.9%	HP1184	conserved hypothetical integral membrane protein	23.5%	HP1149	conserved hypothetical protein	24.7%
Other			HP1185	conserved hypothetical integral membrane protein	55.5%	HP1160	conserved hypothetical protein	34.7%
HP0876	iron-regulated outer membrane protein (frpB)	27.6%	HP1225	conserved hypothetical integral membrane protein	31.6%	HP1162	conserved hypothetical protein	34.8%
HP0915	iron-regulated outer membrane protein (frpB)	28.1%	HP1235	conserved hypothetical integral membrane protein	29.0%	HP1214	conserved hypothetical protein	21.5%
HP0916	iron-regulated outer membrane protein (frpB)	28.8%	HP1235	conserved hypothetical integral membrane protein	30.9%	HP1221	conserved hypothetical protein	42.4%
HP1129	biopolymer transport protein (exbD)	29.7%	HP1330	conserved hypothetical integral membrane protein	41.7%	HP1240	conserved hypothetical protein	22.5%
HP1130	biopolymer transport protein (exbD)	33.5%	HP1331	conserved hypothetical integral membrane protein	33.6%	HP1242	conserved hypothetical protein	42.3%
HP1339	biopolymer transport protein (exbB)	46.8%	HP1343	conserved hypothetical integral membrane protein	49.1%	HP1259	conserved hypothetical protein	44.6%
HP1340	biopolymer transport protein (exbB)	35.8%	HP1363	conserved hypothetical integral membrane protein	33.1%	HP1284	conserved hypothetical protein	36.8%
HP1445	biopolymer transport protein (exbB)	45.5%	HP1407	conserved hypothetical integral membrane protein	22.4%	HP1291	conserved hypothetical protein	26.3%
HP1446	biopolymer transport protein (exbD)	36.2%	HP1466	conserved hypothetical integral membrane protein	30.9%	HP1335	conserved hypothetical protein	33.3%
HP1512	iron-regulated outer membrane protein (frpB)	26.6%	HP1484	conserved hypothetical integral membrane protein	41.2%	HP1337	conserved hypothetical protein	27.2%
HP0653	nonheme iron-containing ferritin (pfr)	99.4%	HP1509	conserved hypothetical integral membrane protein	34.3%	HP1338	conserved hypothetical protein	36.2%
HP1341	siderophore-mediated iron transport protein (tonB)	37.2%	HP1548	conserved hypothetical integral membrane protein	30.6%	HP1394	conserved hypothetical protein	33.6%
OTHER CATEGORIES			HP1551	conserved hypothetical integral membrane protein	38.8%	HP1401	conserved hypothetical protein	27.5%
General			HP1558	conserved hypothetical integral membrane protein	30.9%	HP1413	conserved hypothetical protein	41.6%
HP0924	4-oxalocrotonate tautomerase (dmpl)	37.7%	HP1588	conserved hypothetical integral membrane protein	35.1%	HP1414	conserved hypothetical protein	27.4%
HP1034	ATP-binding protein (yixH)	36.3%	HP1589	conserved hypothetical integral membrane protein	35.1%	HP1417	conserved hypothetical protein	23.7%
HP1000	PARA	29.7%	HP1600	conserved hypothetical integral membrane protein	48.2%	HP1423	conserved hypothetical protein	40.3%
HP1139	SpoQJ regulator (soj)	47.4%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1428	conserved hypothetical protein	37.8%
HP0827	ss-DNA binding protein 12RNP2 precursor	46.5%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1443	conserved hypothetical protein	37.9%
Adaptations and atypical conditions			HP1609	conserved hypothetical integral membrane protein	31.4%	HP1449	conserved hypothetical protein	39.0%
HP1496	general stress protein (ctc)	26.5%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1453	conserved hypothetical protein	26.8%
HP1483	gerC2 protein (gerC2)	33.3%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1459	conserved hypothetical protein	30.1%
HP0927	heat shock protein (hspX)	32.8%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1504	conserved hypothetical protein	23.9%
HP0280	heat shock protein B (hspB)	27.2%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1510	conserved hypothetical protein	30.6%
HP1228	invasion protein (invA)	38.2%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1533	conserved hypothetical protein	25.4%
HP0970	nickel-cobalt-cadmium resistance protein (nccB)	21.1%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1570	conserved hypothetical protein	40.5%
HP1444	small protein (smpB)	42.1%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1573	conserved hypothetical protein	42.2%
HP0930	stationary-phase survival protein (surE)	37.7%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1587	conserved hypothetical protein	39.0%
HP0315	virulence associated protein D (vapD)	70.2%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1588	conserved hypothetical protein	32.0%
HP0967	virulence associated protein D (vapD)	28.9%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1589	conserved hypothetical protein	35.1%
HP1248	virulence associated protein homolog (vacB)	36.0%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1713	conserved hypothetical protein	41.8%
HP0885	virulence factor mviN protein (mviN)	33.5%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0028	conserved hypothetical secreted protein	42.1%
Colicin-related functions			HP1609	conserved hypothetical integral membrane protein	31.4%	HP0139	conserved hypothetical secreted protein	37.1%
HP1126	colicin tolerance-like protein (tolB)	25.7%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0160	conserved hypothetical secreted protein	30.8%
HP0428	phage/colicin/tellurite resistance cluster terY protein	25.6%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0190	conserved hypothetical secreted protein	31.4%
Drug and analog sensitivity			HP1609	conserved hypothetical integral membrane protein	31.4%	HP0211	conserved hypothetical secreted protein	24.3%
HP1431	16S rRNA (adenosine-N6,N6)-dimethyl-transferase (ksa)	35.5%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0235	conserved hypothetical secreted protein	31.5%
HP0606	membrane fusion protein (mtrC)	24.2%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0257	conserved hypothetical secreted protein	29.2%
HP0830	modulator of drug activity (mda68)	62.3%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0320	conserved hypothetical secreted protein	36.4%
HP1476	phenylacrylic acid decarboxylase	39.7%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0506	conserved hypothetical secreted protein	29.8%
HP1165	tetracycline resistance protein tetA(P), putative	27.0%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0518	conserved hypothetical secreted protein	96.9%
Transposon-related functions			HP1609	conserved hypothetical integral membrane protein	31.4%	HP0785	conserved hypothetical secreted protein	26.8%
HP1008	IS200 insertion sequence from SARA17	33.9%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0949	conserved hypothetical secreted protein	39.7%
HP0414	IS200 insertion sequence from SARA17	33.9%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0977	conserved hypothetical secreted protein	29.4%
HP0988	IS605 transposase (tnpA)	97.2%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0980	conserved hypothetical secreted protein	57.4%
HP0998	IS605 transposase (tnpA)	97.2%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1075	conserved hypothetical secreted protein	42.9%
HP1006	IS605 transposase (tnpA)	97.2%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1098	conserved hypothetical secreted protein	27.0%
HP1535	IS605 transposase (tnpA)	97.2%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1117	conserved hypothetical secreted protein	32.3%
HP0437	IS605 transposase (tnpA)	97.2%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1216	conserved hypothetical secreted protein	31.9%
HP0989	IS605 transposase (tnpB)	93.4%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1285	conserved hypothetical secreted protein	38.0%
HP0997	IS605 transposase (tnpB)	93.4%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1286	conserved hypothetical secreted protein	37.5%
HP1095	IS605 transposase (tnpB)	93.4%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1484	conserved hypothetical secreted protein	27.4%
HP1534	IS605 transposase (tnpB)	93.4%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1488	conserved hypothetical secreted protein	29.8%
HP0438	transposase-like protein, PS3IS	33.8%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1551	conserved hypothetical secreted protein	42.7%
HP1007	transposase-like protein, PS3IS	34.3%	HP1609	conserved hypothetical integral membrane protein	31.4%	UNKNOWN		
Other			HP1609	conserved hypothetical integral membrane protein	31.4%	General		
HP0739	2-hydroxy-6-oxohepta-2,4-dienoate hydrolase	30.1%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0390	adhesin-thiol peroxidase (tagD)	38.3%
HYPOTHETICAL			HP1609	conserved hypothetical integral membrane protein	31.4%	HP1193	aldo-keto reductase, putative	46.6%
General			HP1609	conserved hypothetical integral membrane protein	31.4%	HP0872	alkylphosphonate uptake protein (phnA)	61.1%
HP0831	conserved hypothetical ATP binding protein	32.3%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0207	ATP-binding protein (mpr)	39.9%
HP0066	conserved hypothetical ATP-binding protein	34.7%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0136	bacterioferritin coregulatory protein (bcp)	35.5%
HP0269	conserved hypothetical ATP-binding protein	37.7%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0485	catalase-like protein	30.8%
HP0312	conserved hypothetical ATP-binding protein	34.1%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1104	cinnamyl-alcohol dehydrogenase	44.0%
HP1321	conserved hypothetical ATP-binding protein	30.8%	HP1609	conserved hypothetical integral membrane protein	31.4%	EL3-2 (cad)		
HP1430	conserved hypothetical ATP-binding protein	38.1%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0981	exonuclease VII-like protein (xseA)	42.5%
HP1507	conserved hypothetical ATP-binding protein	51.6%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0669	GTP-binding protein (gtpI)	48.1%
HP1567	conserved hypothetical ATP-binding protein	40.9%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0303	GTP-binding protein (obg)	48.2%
HP1026	conserved hypothetical helicase-like protein	35.2%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0634	GTP-binding protein homologue (yphC)	36.7%
HP0022	conserved hypothetical integral membrane protein	30.8%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0480	GTP-binding protein, fusa-homolog (yihK)	54.1%
HP0189	conserved hypothetical integral membrane protein	43.1%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP1489	lipase-like protein	21.7%
HP0226	conserved hypothetical integral membrane protein	27.6%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0405	nifS-like protein	27.3%
HP0228	conserved hypothetical integral membrane protein	43.2%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0221	nifU-like protein	37.3%
HP0234	conserved hypothetical integral membrane protein	32.4%	HP1609	conserved hypothetical integral membrane protein	31.4%	HP0658	PET112-like protein	45.4%
			HP1609	conserved hypothetical integral membrane protein	31.4%	HP0089	pfs protein (pfs)	34.5%
			HP1609	conserved hypothetical integral membrane protein	31.4%	HP0322	poly E-rich protein	28.7%
			HP1609	conserved hypothetical integral membrane protein	31.4%	HP0625	protein E (gpcP)	47.7%
			HP1609	conserved hypothetical integral membrane protein	31.4%	HP0431	protein phosphatase 2C homolog (ptc1)	30.7%
			HP1609	conserved hypothetical integral membrane protein	31.4%	HP0624	solute-binding signature and mitochondrial signature protein (aspB)	26.4%
			HP1609	conserved hypothetical integral membrane protein	31.4%	HP0377	thiol:sulfide interchange protein (dsbC), putative	26.4%

A cytokine-responsive I κ B kinase that activates the transcription factor NF- κ B

Joseph A. DiDonato*, Makio Hayakawa*, David M. Rothwarf*, Ebrahim Zandi* & Michael Karin

Laboratory of Gene Regulation and Signal Transduction, Department of Pharmacology, University of California at San Diego, 9500 Gilman Drive, La Jolla, California 92093-0636, USA

* These authors contributed equally to this work.

Nuclear transcription factors of the NF- κ B/Rel family are inhibited by I κ B proteins, which inactivate NF- κ B by trapping it in the cell cytoplasm. Phosphorylation of I κ Bs marks them out for destruction, thereby relieving their inhibitory effect on NF- κ B. A cytokine-activated protein kinase complex, IKK (for I κ B kinase), has now been purified that phosphorylates I κ Bs on the sites that trigger their degradation. A component of IKK was molecularly cloned and identified as a serine kinase. IKK turns out to be the long-sought-after protein kinase that mediates the critical regulatory step in NF- κ B activation.

Transcription factors of the NF- κ B/Rel family are regulated through interactions with I κ Bs, which are inhibitory proteins^{1,2}. With the exception of mature B cells where they are constitutively nuclear, in all other cell types NF- κ B dimers are kept in the cytoplasm through association with the I κ Bs, which mask their nuclear localization sequence^{1,2}. In response to diverse extracellular stimuli, including proinflammatory cytokines, viral infection, oxidants, phorbol esters and ultraviolet irradiation, the I κ Bs are rapidly phosphorylated at two serines within their amino-terminal regulatory domains. In two I κ B proteins, site-directed mutagenesis^{3–7} and biochemical mapping⁷ has identified these serines as S32 and S36 in I κ B α and S19 and S23 in I κ B β . Phosphorylation at these sites triggers polyubiquitination of the I κ Bs^{7–9} and, as with other proteins¹⁰, targets them for rapid degradation by the 26S proteasome^{7,8}. Proteasome inhibitors cause accumulation of I κ B, which in stimulated cells is N-terminally phosphorylated, and inhibit NF- κ B activation^{11,12}. Substitution of two lysines in I κ B α , K20 and K21, one of which is conserved in I κ B β , with arginine residues decreases the rate of inducible I κ B α ubiquitination and degradation without affecting its phosphorylation⁷.

As N-terminal phosphorylation of the I κ Bs occurs before and is necessary for their ubiquitination and degradation, it is likely to be the principal point of control through which diverse stimuli effect NF- κ B activation¹³. However, it is not yet clear whether increased N-terminal phosphorylation of the I κ Bs is due to activation of an I κ B kinase or inhibition of an I κ B phosphatase^{14,15}. Part of this doubt is due to the fact that a cytokine-responsive protein kinase(s) phosphorylating I κ B α at S32/36 or I κ B β at S19/23 has not been previously identified. In addition, okadaic acid, an inhibitor of protein-phosphatase 2A (PP2A), is a potent NF- κ B activator and an inducer of I κ B N-terminal phosphorylation^{8,14,15}. In fact, an activity that phosphorylates I κ B α at S32/36 was detected in okadaic acid-treated cell extracts and had an apparent relative molecular mass (M_r) of 700K (ref. 15). It was not established, however, whether this activity is stimulated by proinflammatory cytokines or other agonists and, if so, whether its activation kinetics correlate with those of I κ B phosphorylation.

To understand better the signalling mechanism responsible for NF- κ B activation, we tested cell extracts for the presence of a protein kinase activity that is activated by tumour-necrosis factor (TNF) and phosphorylates I κ B α at S32/36. Such an activity was detected in extracts of TNF-treated HeLa cells and its substrate specificity

and kinetics of activation correlated well with those of I κ B α phosphorylation in living cells⁷. We purified this activity and determined a partial peptide sequence for one of its components. Molecular cloning and functional analysis identified this subunit as a protein kinase whose associated I κ B kinase activity is rapidly stimulated by proinflammatory cytokines. This activity is inhibited upon dephosphorylation with PP2A. We demonstrate that this protein kinase, IKK α , is critical for NF- κ B activation in response to proinflammatory cytokines.

Purification of an I κ B α kinase

To investigate whether phosphorylation of I κ B α at S32/36 in response to TNF is due to activation of an I κ B kinase specific for these sites, we fractionated extracts of non-stimulated and TNF-stimulated HeLa cells on a Mono-Q column. Column fractions were tested for their ability to phosphorylate a glutathione S-transferase (GST)–I κ B α fusion protein containing the N-terminal regulatory domain (residues 1 to 54). The specificity of the kinase and its physiological relevance were examined by using the mutant substrate GST–I κ B α (1–54; TT), in which serines 32 and 36 are substituted with threonines⁷. In intact cells this substitution decreases the efficiency of I κ B α phosphorylation, indicating that the relevant kinase strongly prefers serines over threonines⁷. As shown in Fig. 1a, TNF stimulation of HeLa cells resulted in rapid activation of a kinase activity that phosphorylates GST–I κ B α (1–54). This activity did not phosphorylate GST–I κ B α (1–54; TT). Fractionation of extracts of non-stimulated and TNF-stimulated HeLa cells by gel filtration revealed a TNF-stimulated I κ B α (1–54) kinase activity that eluted with an apparent M_r of ~900K (Fig. 1b). This activity was not detected when GST–I κ B α (1–54; TT) was used as a substrate and also did not phosphorylate GST–I κ B α (1–54; AA), in which serines 32 and 36 were substituted with alanines (Fig. 1c). This kinase activity was activated rapidly, peaking 5–10 min after TNF addition (data not shown).

We purified this kinase activity from large-scale cultures of HeLa cells stimulated with TNF. Cytosolic extracts (S-100 fraction) were chromatographed on Q-Sepharose and the active fractions were pooled and chromatographed through a γ -phosphate-linked ATP-Sepharose¹⁶ affinity column. Bound proteins were eluted with ATP and the active fractions pooled and chromatographed on a Superose-6-gel-filtration column. This procedure resulted in ~25,000-fold enrichment of I κ B α kinase activity (Table 1). The purified

kinase, which we call I κ B kinase (IKK), has the same substrate specificity as the TNF-stimulated activity already described (Fig. 2c). To examine the polypeptide composition of IKK, we analysed the Superose-6 column fractions by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and silver staining. Only two polypeptides, of 85K and 87K, clearly co-eluted with IKK activity at this stage (Fig. 2a). As even the most active fraction contained a number of polypeptides, we applied the pooled Superose-6 IKK fractions to an affinity column of a GST–I κ B α (1–54; AA)_{8x} fusion protein covalently linked to agarose. After extensive washing, the column was developed with a NaCl gradient. This resulted in a further fourfold enrichment of IKK activity (Table 1). The composition of the material retained on the affinity column is shown in Fig. 2b. Although this fraction contains several polypeptides, we consistently detected enrichment of the 85K and 87K band, as well as of the 64K band. It is not clear yet whether the other bands represent additional IKK components or contaminants. Gel-filtration analysis of this material indicated that it continued to elute as a very large (~900K) complex.

Cloning of an IKK protein-kinase subunit

We purified the 85K polypeptide associated with IKK to homogeneity by preparative SDS–PAGE and subjected it to microsequen-

Table 1 Purification of IKK activity

Step	Protein yield (mg)	Total IKK activity*
S-100	900	–
Q-Sepharose	40	100%
ATP-Sepharose	0.5	170%
Superose 6	0.035	230%
I κ B α -affinity	0.009	160%

The yields and relative activities for the various steps involved in IKK purification are shown from a typical 15-litre HeLa cell suspension culture.

* The total IKK activity is approximate because of the difficulty in estimating IKK activity in the starting material owing to the presence of other kinase activities. Also, owing to removal of inhibitors, the kinase activity increases instead of decreases during most of the steps.

cing. In addition to using oligonucleotide primers based on the obtained sequence, we searched Genbank for related sequences and found that both peptides were contained within a partial sequence of a putative human serine/threonine kinase with unknown function named CHUK¹⁷. Polymerase chain reaction (PCR) and library screening were used to isolate the complete complementary DNA. Nucleotide sequencing revealed the presence of an open reading frame for a 744-amino-acid polypeptide (Fig. 3; Genbank accession number AF009225). The N-terminal half of the predicted polypeptide contains a 301-amino-acid protein-kinase domain,

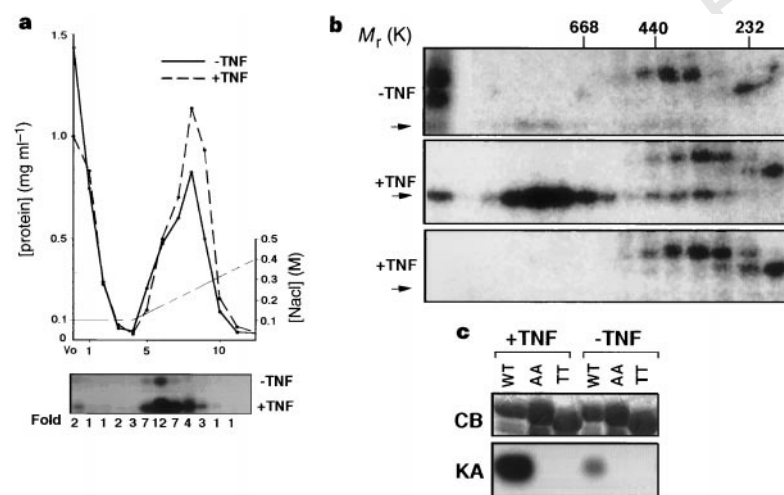
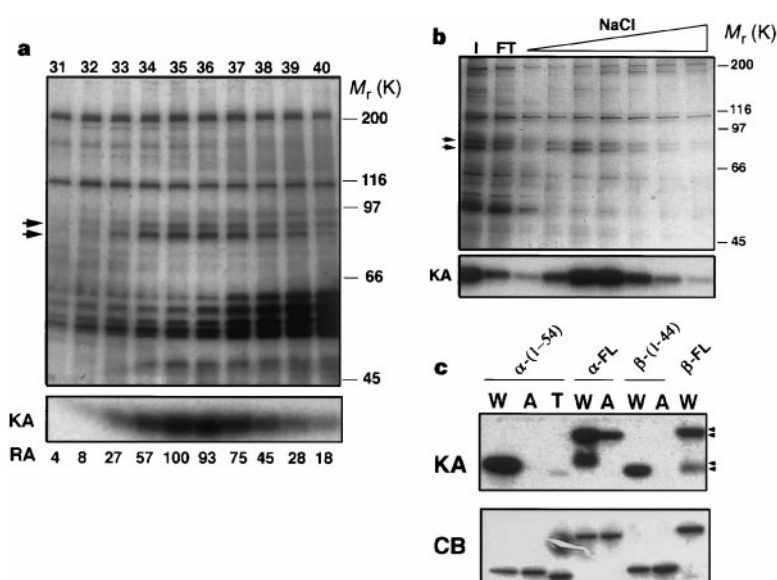


Figure 2 Purification and substrate specificity of I κ B α kinase (IKK). **a**, Cytoplasmic extracts from TNF-stimulated HeLa cells were purified as described. Top, active Superose-6 fractions analysed by SDS–PAGE and silver staining. Fraction numbers are indicated at the top, M_r markers on the right. Arrows, bands that coelute with kinase activity. Bottom, kinase assay (KA) of each fraction with relative activity (RA) indicated underneath. **b**, Fraction 35 was further purified by substrate-affinity chromatography using GST–I κ B α (1–54; AA)_{8x} covalently bound to Sepharose. I, input; FT, flow through; NaCl, eluate. The 85K and 87K bands are indicated. Bottom, kinase assay (KA). **c**, Purified IKK was tested for its ability to phosphorylate GST–I κ B α (1–54)(WT), GST–I κ B α (1–54; AA), GST–I κ B α (1–54; TT), full-length I κ B α (α -FL), GST–I κ B β (1–44)(WT), GST–I κ B β (1–44; AA), and full-length I κ B β (β -FL). Kinase assay (KA, top) and Coomassie blue (CB, bottom)-stained proteins are shown.

Figure 1 Identification of a TNF-stimulated I κ B α kinase activity. **a**, Extracts of non-stimulated or TNF-stimulated HeLa cells were fractionated on a Mono-Q column. Top, elution profile of bulk protein. Bottom, elution of I κ B α kinase activity assayed with GST–I κ B α (1–54) as a substrate. Fold-stimulation of kinase activity in each fraction was determined by phosphorimaging and is indicated at the bottom. **b**, Extracts of TNF-stimulated and non-stimulated HeLa cells fractionated on a Superose-6 column. Top, phosphorylation of GST–I κ B α (1–54) (arrow) by non-stimulated and TNF-stimulated extracts. Bottom, phosphorylation of GST–I κ B α (1–54; TT) by TNF-stimulated cell extract. Positions of M_r markers are indicated at the top; the first lane in each panel is input. **c**, Substrate specificity of the most active fractions of each extract. GST–I κ B α (1–54) (WT, GST–wild type)/GST–I κ B α (1–54; AA), and GST–I κ B α (1–54; TT) were used as substrates. Coomassie blue (CB)-stained substrates and ³²P-labelling (KA, kinase assay) are shown.



whereas its C-terminal half, as previously reported¹⁷, contains several protein-interaction motifs, including a leucine zipper and a helix–turn–helix motif. In the light of its function, we renamed the protein encoded by this cDNA as IKK α (for α subunit of IKK).

Cell-free translation of *in vitro* generated IKK α transcripts resulted in production of a polypeptide of the expected size (Fig. 4a; cell-free-translated IKK α migrates more slowly than the p85 IKK subunit owing to the presence of a haemagglutinin(HA) epitope). We tested this protein for I κ B α kinase activity. HA–IKK α protein was immunoprecipitated with HA antibody, extensively washed in urea at concentrations of up to 3 M, and the immune complexes incubated with various GST–I κ B proteins or GST–c-Jun in the presence of [γ -³²]ATP. This resulted in the efficient phosphorylation of either GST–I κ B α (1–54) or full-length GST–I κ B α (Fig. 4b). Phosphorylation of either GST–I κ B β (1–44) or full-length GST–I κ B β was less efficient, and GST–I κ B α (1–54; TT), GST–I κ B α (1–54; AA), GST–I κ B β (1–44; AA) or GST–c-Jun(1–79) were not phosphorylated. This profile of substrate specificity is identical to that of native IKK purified from TNF-treated HeLa cells (Fig. 2c).

Expression and activation of IKK α

The most efficient NF- κ B activators are the proinflammatory cytokines interleukin (IL)-1 and TNF^{13,18}. These cytokines are also

efficient inducers of I κ B α phosphorylation and degradation⁷. We examined the effect of TNF and IL-1 on IKK α -associated kinase activity. HA–IKK α was transiently expressed in HeLa cells. As a control, we transiently transfected HeLa cells with a HA–JNK1 expression vector. After 36 h, these cells were stimulated with TNF or IL-1 for 5 min and the HA-tagged protein kinases immunoprecipitated. HA–IKK α immunoprecipitates from TNF- or IL-1-stimulated cells phosphorylated GST–I κ B α (1–54) but did not phosphorylate GST–c-Jun(1–79) (Fig. 5a). By contrast, HA–JNK1 phosphorylated GST–c-Jun(1–79) but not GST–I κ B α (1–54). The extent of IKK α activation by TNF or IL-1 was similar to the extent of JNK activation by these cytokines. Time-course analysis indicated that stimulation of IKK α activity occurred rapidly (Fig. 5b). Maximal I κ B α kinase activity was reached within 5 min of TNF or IL-1 addition. After 15–30 min, this activity rapidly declined. IKK α activity was also moderately stimulated by the phorbol ester 12-*O*-tetradecanoylphorbol-13-acetate acetate (TPA; data not shown).

We investigated the substrate specificity of HA–IKK α immune complexes isolated from TNF-stimulated cells and found it to be identical to that of purified IKK (data not shown). Phosphopeptide mapping on Tris–Tricine gels⁷ revealed that phosphorylation of GST–I κ B α (1–54) occurred on the peptide that contains S32 and



Primers:

1: 5'-CCCCATATGTACCAGCATCGGGAA

2: 5'-ACIAT^GTT^TTCIGG^TCTT

3: 5'-CCCCTCGAGTTCTGTAAACCACT

Figure 3 The predicted amino-acid sequence of IKK α : the two peptides whose sequence was determined are overlined. The kinase domain (KD), leucine zipper (LZ) and helix–loop–helix (HLH) are indicated. Peptide sequences used for primer design are indicated by underlining arrows; the sequences of the three primers are shown.

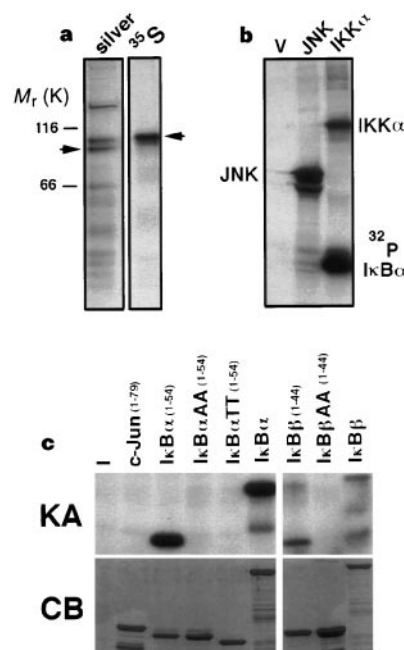


Figure 4 Cell-free translation of IKK α generates I κ B kinase activity. **a**, Cell-free-translated HA–IKK α labelled with ³⁵S-methionine was separated by SDS–PAGE next to affinity-purified IKK stained with silver. **b**, Mock-programmed (V), HA–IKK α - or HA–JNK1-programmed reticulocyte lysates were immunoprecipitated with HA antibody and, after washing with 3M urea, immune complex kinase activity was assayed with GST–I κ B α (1–54) as substrate. ³⁵S- and ³²P-labelled proteins were detected by autoradiography. **c**, After translation of HA–IKK α and immunoprecipitation, the immune complexes were incubated with either no exogenous substrate (–) or the indicated substrates in kinase buffer and [γ -³²P]ATP. Top, kinase assay (KA) revealed by autoradiography; bottom, Coomassie blue (CB)-stained substrates.

S36 (Fig. 6). The alanine-substitution mutant GST-I κ B α (1–54; AA) was not phosphorylated at all (data not shown; but see Figs 2c and 4c). The use of single-substitution mutants, I κ B α (A32) or I κ B α (A36), indicated that both purified IKK and HA-IKK α immune complexes phosphorylated both S32 and S36 with comparable efficiency (Fig. 6b, c). However, phosphopeptide mapping of full-length I κ B α phosphorylated by either purified IKK or HA-IKK α immune complexes revealed that, in addition to the peptide containing S32 and S36, a smaller amount of 32 P was incorporated into a peptide spanning amino acids 264 to 314 (Fig. 6a, c). When

full-length I κ B α (A32/36) was used as a substrate, all of the 32 P was incorporated into that C-terminal peptide. Phosphoaminoacid analysis indicated that phosphorylation occurred at serine residues. Previously we have localized the constitutive phosphorylation sites of I κ B α to the same peptide near its C terminus⁷ (J.A.D., unpublished results). It is not yet clear, however, whether IKK phosphorylates exactly the same sites.

IKK α is involved in NF- κ B activation

To investigate further the role of IKK α in I κ B phosphorylation in

Figure 5 Expression and cytokine activation of IKK in mammalian cells. **a**, HeLa cells were transiently transfected with either HA-JNK1 or HA-IKK α expression vectors. After 36 h, cells were incubated with either control medium (–), TNF or IL-1 for 5 min and the activities of the HA-tagged kinases were determined by immune-complex kinase assays with GST-c-Jun(1–79) or GST-I κ B α (1–54) as substrates. The recovery of HA-JNK1 and HA-IKK α was determined by immunoblotting. **b**, HeLa cells were transfected with HA-IKK α expression vector and were stimulated after 36 h with either TNF or IL-1 for the indicated times, at which point cells were lysed and IKK activity determined by immune-complex kinase assay.

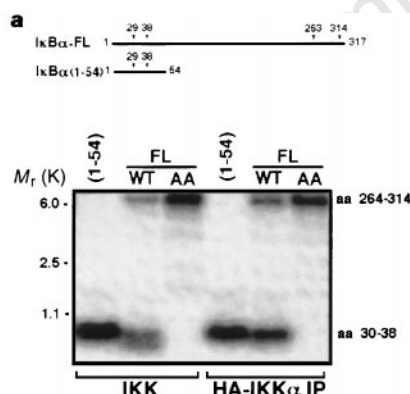
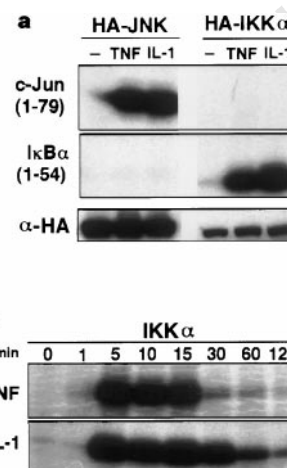
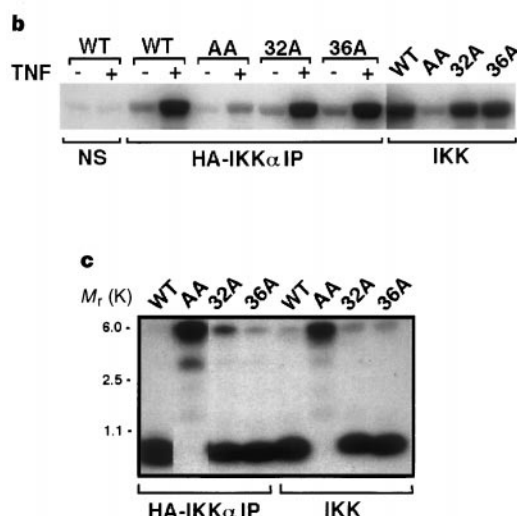


Figure 6 Mapping of IKK phosphorylation sites on I κ B α . **a**, GST-I κ B α (1–54), full-length (FL) WT or A32/36 I κ B α were phosphorylated by either purified IKK or HA-IKK α immune complexes isolated from TNF-stimulated cells. Equal counts of each protein were digested with trypsin and the resulting digests separated on Tris-Tricine gels together with M_r markers. The $\sim$ 0.9K peptide corresponds to amino acids 30–38 and the $\sim$ 6K peptide corresponds to amino acids 263–314 (ref. 7). The location of these tryptic peptides is shown at the top. **b**, Extracts were prepared from HeLa cells stably expressing HA-IKK α that were either non-stimulated or stimulated with TNF for 10 min and immunoprecipitated with either nonspecific (NS) or HA-antibody. The immune complexes were used to phosphorylate full-length WT I κ B α , I κ B α (A32/36), I κ B α (A32) or I κ B α (A36). The same proteins were also phosphorylated by purified IKK. Equal amounts of each substrate were analysed by SDS-PAGE and autoradiography. **c**, The proteins in **b** were subjected to tryptic phosphopeptide mapping and the digests resolved on a Tris-Tricine gel.



intact cells, we generated a stably transfected pool of HT-29 cells expressing HA-IKK α . Previous studies indicated that TNF induces partial I κ B α degradation in these cells with slower kinetics than in other cell lines, especially HeLa (J.A.D., unpublished results). Expression of HA-IKK α in HT-29 cells markedly accelerated I κ B α phosphorylation (indicated by the appearance of a slower migrating form⁷) and degradation in response to TNF (Fig. 7). The kinetics of I κ B α degradation correlated with the kinetics of IKK activation.

We tested the effect of raised or reduced IKK α expression on NF- κ B activation by transient transfection. A 2 $\times$ NF- κ B-Luc

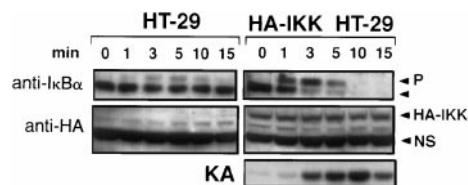


Figure 7 Expression of IKK α accelerates I κ B α degradation. Pools of HT-29 cells stably expressing HA-IKK α and the parental cells (HT-29) were stimulated with TNF for the indicated times at which point the cells were collected and lysed. Cell lysates were tested by immunoblotting for I κ B α degradation⁷ and HA-IKK α expression, as well as IKK activity of HA-IKK α immune complexes (KA). The basal and the hyperphosphorylated forms of I κ B α are indicated, as is the HA-IKK α band. NS, nonspecific. HA immunoprecipitates of the parental cells do not contain IKK activity (not shown).

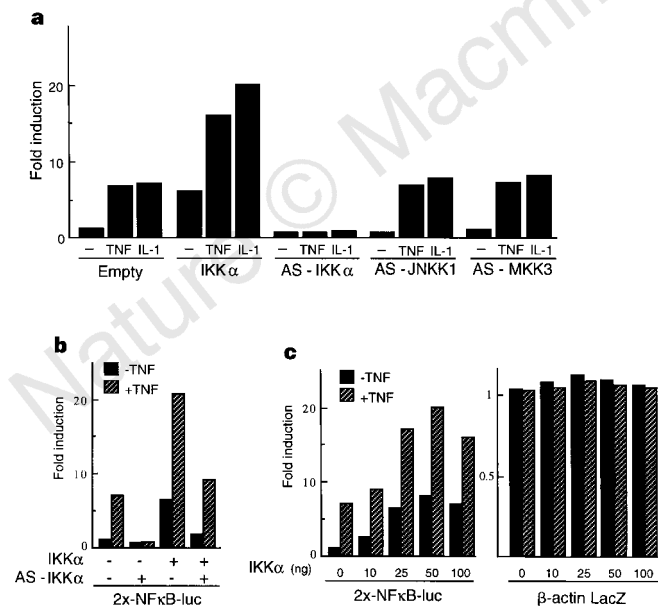


Figure 8 IKK α and NF- κ B activation. **a**, HeLa cells were transiently cotransfected with 2 $\times$ NF- κ B-Luc reporter¹² together with empty expression vector, HA-IKK α expression vector, antisense (AS) IKK α vector (IKK α cDNA cloned in 3' to 5' orientation downstream of the β -actin promoter), antisense JNKK1 vector or antisense MKK3 vector. After 36 h, cultures were divided and either left untreated or incubated with either TNF or IL-1. After 3 h, luciferase activity was determined. The average fold-induction of luciferase activity in 3 separate experiments is shown. **b**, Cells were cotransfected with empty expression vector, HA-IKK α or antisense IKK α vectors, alone or in combination. After 36 h, cells were either stimulated or not with TNF for 3 h and luciferase activity determined. **c**, Cells were cotransfected with the indicated reporters together with the indicated amounts of HA-IKK α expression vector. After 36 h, cells were stimulated or not with TNF; after 3 h, luciferase and β -galactosidase activities were determined. Results are also shown from a β -actin-Lac Z reporter.

reporter¹² was cotransfected with an empty expression vector, an HA-IKK α expression vector, or an antisense (AS) IKK α vector into HeLa cells. Expression of HA-IKK α potentiated basal reporter gene activity and its stimulation by TNF or IL-1 (Fig. 8a). By contrast, cotransfection with AS-IKK α blocked induction of 2 $\times$ NF- κ B reporter activity by IL-1 or TNF. In other experiments, AS-IKK α inhibited induction of 2 $\times$ NF- κ B-Luc by TPA (data not shown). The inhibition was specific as it was not seen upon cotransfection of either JNKK1 or MKK3 antisense vectors. In addition, the inhibitory effect of antisense IKK α was reversed upon cotransfection of the HA-IKK α vector (Fig. 8b). Furthermore, immunoblot analysis indicated that AS-IKK α inhibited expression of HA-IKK α but not HA-JNKK1 (data not shown). The stimulatory effect of HA-IKK α on NF- κ B reporter activity was specific as it was not detected using a β -actin-LacZ reporter (Fig. 8c).

IKK is inactivated by PP2A

The identification of a cytokine-activated I κ B kinase suggests that the activation of NF- κ B and induction of I κ B α phosphorylation by the PP2A-inhibitor okadaic acid^{14,15} could be due to negative regulation of IKK itself by PP2A. Such an interpretation is consistent with the rapid inactivation of IKK activity following its stimulation by TNF or IL-1. To test this possibility, we performed the experiment shown in Fig. 9a. Whereas okadaic acid did not affect I κ B α phosphorylation by IKK purified from TNF-treated HeLa cells, inclusion of PP2A in the reaction or preincubation of IKK with PP2A diminished I κ B α phosphorylation. This effect was inhibited by okadaic acid. To determine the target for PP2A action, we preincubated purified IKK enzyme with PP2A and then performed the kinase reaction in the presence of okadaic acid, which caused a large decrease in I κ B α phosphorylation. When PP2A was added during the kinase reaction together with okadaic acid, no effect on I κ B α phosphorylation was seen. Hence, IKK activity is sensitive to PP2A. We also found that, once phosphorylated by IKK, I κ B α was not dephosphorylated by PP2A (Fig. 9b).

The I κ B kinase activity associated with IKK α is also sensitive to PP2A. HA-IKK α immune complexes from transiently transfected non-stimulated or TNF-stimulated cells were treated or not with PP2A and tested for I κ B α phosphorylation in the absence or presence of okadaic acid. Like the purified enzyme, the activity of HA-IKK α was sensitive to PP2A (Fig. 9c). Consistent with these results, we find that incubation of HeLa or HT-29 cells, stably expressing HA-IKK α , with okadaic acid caused a slow but considerable increase in IKK activity (Fig. 9d). In addition, cotransfection with the AS-IKK α vector inhibited induction of 2 $\times$ NF κ B-Luc by okadaic acid (data not shown).

Discussion

The proinflammatory cytokines TNF and IL-1 exert many of their effects through activation of transcription factors AP-1 and NF- κ B, which mediate induction of genes encoding other cytokines and various proteins involved in the inflammatory response^{18–20}. We understand the immediate events involved in TNF (refs 21–23) and IL-1 (refs 24, 25) signalling and how they lead to activation of the Jun N-terminal kinase (JNK) and p38 MAPK cascades that mediate induction of AP-1 activity²⁶ (Z. G. Liu, unpublished results), but not how these early signalling events lead to NF- κ B activation. Although I κ B phosphorylation is critical for cytokine-mediated NF- κ B activation, the protein kinase(s) responsible was previously unidentified. It was even suggested that increased I κ B phosphorylation following cytokine treatment was due to inhibition of a phosphatase. We have now described the purification and molecular cloning of a component of the cytokine-activated protein kinase complex IKK which is responsible for inducible I κ B phosphorylation.

Several lines of evidence indicate that IKK is the critical protein kinase mediating NF- κ B activation by TNF or IL-1. First, IKK activity is rapidly stimulated by TNF, IL-1 or TPA and its kinetics of

activation match those of I κ B α phosphorylation and degradation in intact cells. Second, IKK phosphorylates I κ B α and β on the same residues, S32/36 and S19/23, respectively, that are phosphorylated in intact cells in response to TPA, IL-1 or TNF⁷. Phosphorylation of these serines triggers the polyubiquitination and degradation of I κ B³⁻⁸. Third, IKK does not phosphorylate mutants of I κ B α in which these serines are replaced with threonines. Such mutants are poorly phosphorylated in intact cells and are refractory to cytokine-induced degradation⁷. Fourth, elevated expression of IKK α , the subunit of IKK that we have cloned, accelerates I κ B α degradation and potentiates NF- κ B activation, whereas expression of an antisense IKK α construct blocks NF- κ B activation by TNF or IL-1. Fifth, IKK activity is sensitive to PP2A, whereas the specific PP2A inhibitor, okadaic acid, activates NF- κ B^{14,15} and IKK. On the other hand, PP2A does not dephosphorylate I κ B α previously phosphorylated by IKK.

IKK purifies as a very large and stable species composed of several polypeptides. The exact polypeptide composition of IKK remains to be determined and the protein kinase we have cloned, IKK α , is only one subunit of that complex. Although the size of the IKK complex is similar (but not identical) to that of a previously reported I κ B kinase¹⁵, the two activities may not be related. The activity of the enzyme studied by Chen *et al.*¹⁵ depends on its ubiquitination, but we find no evidence that ubiquitination is required for IKK activation. Rather, the critical modification for IKK activation, based on its sensitivity to PP2A, appears to be phosphorylation. In addition, the previously described I κ B kinase activity appears to be constitutive¹⁵, whereas IKK activity is rapidly stimulated by cytokines. Nevertheless, IKK isolated from unstimulated cells has a basal activity, amounting to 5–10% of that of the activated enzyme. Indeed, overexpression of IKK α can enhance NF- κ B activity even in unstimulated cells. In addition, the upstream inputs that lead to IKK activation are weakly active in unstimulated cells. Inhibition of PP2A, the phosphatase that inactivates IKK, results in slow but considerable IKK activation. The nature of the input responsible for basal IKK activity remains to be identified. It has been shown²⁷ that

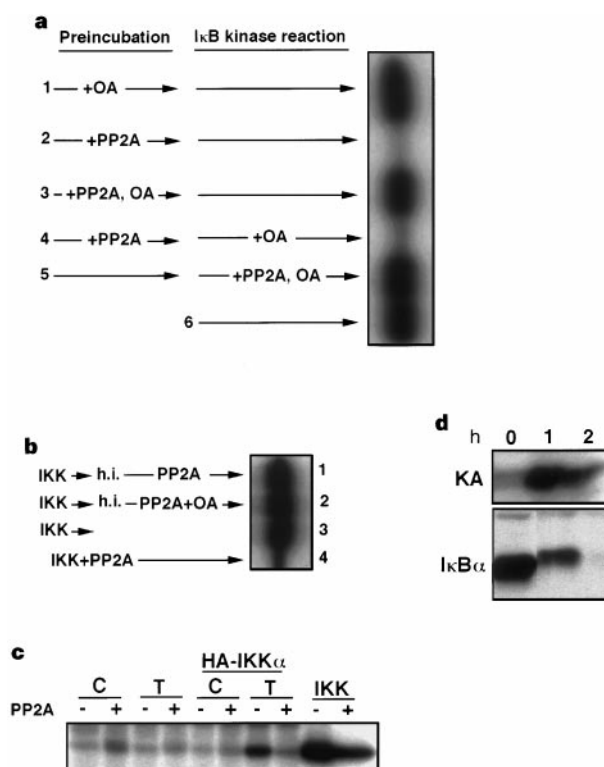
the ubiquitination-dependent I κ B kinase can be further activated *in vitro* by MEKK1; however, expression of a catalytically inactive MEKK1 mutant inhibits JNK activation without affecting NF- κ B activation²⁶ and so far we have been unable to activate IKK *in vitro* with MEKK1 (unpublished results).

As IKK α is a serine/threonine kinase by sequence and its cell-free translation produces an I κ B kinase activity with the same substrate specificity as native IKK, it is likely that it is a catalytic subunit of this complex. IKK α , which is essentially identical in sequence to the putative serine/threonine kinase CHUK, whose function and mechanism of regulation were previously unknown¹⁷, contains protein-interaction motifs in its C-terminal half, including a leucine zipper and a helix–turn–helix. These motifs could mediate interaction of IKK α with other subunits of the IKK complex. Although the I κ Bs are not part of this complex, IKK binds to an I κ B α -affinity column. The IKK subunit responsible for this interaction remains to be identified. Previously, by studying the interaction of JNK with c-Jun, we demonstrated that docking of a signal-regulated protein kinase to its substrate is important for rapid and specific phosphorylation²⁸.

Although in Jurkat cells TNF induces the degradation of I κ B α but not I κ B β ²⁹, in several other cell lines we found that the main difference between the two I κ Bs was in the rate of their degradation: I κ B α is degraded faster than I κ B β ⁷. This could be caused by a limiting amount of a protein kinase with a higher affinity for I κ B α than for I κ B β . IKK exhibits this property *in vitro*, phosphorylating I κ B α more efficiently than I κ B β . Also, if I κ B β degradation contributes to NF- κ B activation during a 3-hour stimulation in HeLa cells, the fact that antisense IKK α RNA blocks NF- κ B activation indicates that IKK may also be an I κ B β kinase. But it remains to be seen whether overexpression or activation of IKK accelerates I κ B β phosphorylation and degradation.

Our results establish IKK as the long-sought-after protein kinase that phosphorylates at least I κ B α in response to proinflammatory cytokines. The identification of IKK and the molecular cloning of one of its subunits will lead to a better understanding of the

Figure 9 IKK is sensitive to PP2A. **a**, IKK was preincubated with or without PP2A catalytic subunit (gift from G. Walter) at 30°C for 1 h and then assayed for GST-I κ B α (1–54) phosphorylation. Where indicated, okadaic acid (OA) was added. **b**, GST-I κ B α (1–54) was phosphorylated by IKK and the kinase was heat-inactivated (h.i.). The phosphorylated substrate was then incubated with PP2A with or without OA as indicated (lanes 1, 2). Lane 3, control kinase reaction; lane 4; PP2A was present throughout the reaction. **c**, HA-IKK α was immunoprecipitated from transiently transfected HeLa cells incubated in either control medium (C) or in the presence of TNF for the last 5 min (T). The immune complexes were preincubated in the absence (–) or present (+) of PP2A for 1 h and then tested for GST-I κ B α (1–54) phosphorylation. As a control, HA immune complexes were also isolated from non-transfected cells (first four lanes). Last two lanes, PP2A, sensitivity of purified IKK. **d**, HeLa cells stably expressing HA-IKK α were incubated with 50 nM okadaic acid for the indicated times, then cells were lysed and the IKK activity assayed with GST-I κ B α (1–54) as substrate (top). Degradation of I κ B α in these cells was determined by immunoblotting (bottom).



signalling pathways originating at the TNF and IL-1 receptors that depend on recruitment of TRAF mediators^{21–25} to elicit activation of NF- κ B and the inflammatory gene-induction response. □

Methods

Protein purification. Suspension cultures of HeLa S3 cells were stimulated with 20 ng ml⁻¹ TNF (R&D Systems) for 5 min at 5×10^6 cells ml⁻¹. Cells were collected and lysed in buffer A (20 mM Tris-Cl, 20 mM NaF, 20 mM β -glycerophosphate, 19 mM PNPP, 500 μ M Na₃VO₄, 2.5 mM metabisulphite, 5 mM benzamidine, 1 mM EDTA, 0.5 mM EGTA, 1 mM PMSF, 10% glycerol, pH 7.6) supplemented with 20 μ g ml⁻¹ aprotinin, 2.5 μ g ml⁻¹ leupeptin, 8.3 μ g ml⁻¹ bestatin, 1.7 μ g ml⁻¹ pepstatin and 0.05% NP-40, by using a glass Dounce homogenizer. After centrifugation at 12,000 r.p.m. for 20 min in a Beckman SS34 rotor, the supernatant was recentrifuged at 38,000 r.p.m. for 80 min in a Beckman Ti 50.1 rotor, all at 4 °C. The supernatant (S-100) was flash-frozen and stored at -80 °C. Thawed S-100 fractions were purified over a 56 ml Q-Sepharose FF column equilibrated with buffer A and 0.1% Brij-35 (buffer B). After washing with buffer B plus 100 mM NaCl, the column was eluted with a linear 100–300 mM NaCl gradient. Fractions containing peak IKK activity were pooled, diluted 1:4 and applied to a 5-ml Hi Trap-Q column (Pharmacia). After washing with buffer B, the column was eluted in buffer B plus 300 mM NaCl. IKK-containing fractions were pooled and diluted 1:1 with ATP-column buffer¹⁶. The material was passed 4 times over a 4-ml γ -phosphate-linked ATP-affinity column¹⁶. After washing with 10 ml ATP column buffer, 0.05% Brij-35, and with 10 ml ATP-column buffer, 0.05% Brij-35, 250 mM NaCl, the column was eluted with ATP-column buffer, 0.05% Brij-35, 250 mM NaCl, 10 mM ATP. After 1:3 dilution with buffer B, the eluate was applied to a 1-ml Hi-Trap-Q column. The column was eluted with buffer B plus 300 mM NaCl. IKK-containing fractions were pooled, concentrated and applied to a Superose-6 column, equilibrated and eluted in the same buffer. IKK-containing fractions were pooled and applied to an affinity column of a GST-I κ B α (1–54; AA)₈ (eight repeats of the N-terminal domain) covalently linked to Sepharose after 1:4 dilution in buffer A. After washing in buffer A the column was eluted with a NaCl gradient.

IKK assay, immunoprecipitation and immunoblotting. Kinase activity was assayed in 20 mM HEPES, 20 mM β -glycerophosphate, 10 mM MgCl₂, 10 mM PNPP, 100 μ M Na₃VO₄, 2 mM DTT, 20 μ M ATP, 10 μ g ml⁻¹ aprotinin, 50–200 mM NaCl, pH 7.5, and (1–10 μ Ci)[γ -³²P]ATP at 30 °C for 30 min. I κ B-substrate proteins were expressed and purified from *E. coli*. HA-IKK α immune complexes were isolated as described³⁰ and washed in kinase buffer containing 3M urea before determining kinase activity. Immunoblotting was done as described⁷.

Peptide sequencing and mapping. Polypeptides to be sequenced were purified by preparative SDS–PAGE. The 85K IKK α band was digested *in situ* with endoprotease Lys-C. The resultant peptides were eluted and purified by reverse-phase chromatography on an ABI 173 microblotter. Their sequence was determined on an ABI Procise Protein MicroSequencer. Phosphopeptide mapping on Tris–Tricine gels and amino-acid analysis were done as described⁷.

Plasmids, cell culture and transfections. The various expression vectors were constructed using standard recombinant DNA procedures. The β -actin promoter³¹ was used to drive expression of both HA-IKK α and antisense IKK α . Transient transfections were done as described^{26,28}. Stably transfected cell lines were generated as described^{7,31} and identified by immunoblot screening of individual clones or pools of clones. Luciferase and β -galactosidase assays have also been described^{12,26}.

Received 1 May; accepted 4 July 1997.

- Beg, A. A. & Baldwin, A. S. Jr The I κ B proteins: multifunctional regulators of Rel/NF- κ B transcription factors. *Genes Dev.* **7**, 2064–2070 (1993).
- Gilmore, T. D. & Morin, P. J. The I κ B proteins: members of a multifunctional family. *Trends Genet.* **9**, 427–433 (1993).
- Brockman, J. A. *et al.* Coupling of a signal response domain in I κ B to multiple pathways for NF- κ B activation. *Mol. Cell. Biol.* **15**, 2809–2818 (1995).
- Brown, K., Gerstberger, F. S., Carlson, L., Franzoso, G. & Siebenlist, U. Control of I κ B α proteolysis by site-specific, signal-induced phosphorylation. *Science* **267**, 1485–1491 (1995).
- Traenkle, E. B. *et al.* Phosphorylation of human I κ B on serines 32 and 36 controls I κ B proteolysis and NF- κ B activation in response to diverse stimuli. *EMBO J.* **14**, 2876–2883 (1995).
- Whiteside, T. *et al.* N- and C-terminal sequences control degradation of MAD3/I κ B α in response to inducers of NF- κ B activity. *Mol. Cell. Biol.* **15**, 5339–5345 (1995).
- DiDonato, J. A. *et al.* Mapping of the inducible I κ B phosphorylation sites that signal its ubiquitination and degradation. *Mol. Cell. Biol.* **16**, 1295–1304 (1996).
- Chen, Z. *et al.* Signal-induced site-specific phosphorylation targets I κ B to the ubiquitin-proteasome pathway. *Genes Dev.* **9**, 1586–1597 (1995).
- Scherer, D. C., Brockman, J., Chen, Z., Maniatis, T. & Ballard, D. Signal-induced degradation of I κ B requires site-specific ubiquitination. *Proc. Natl Acad. Sci. USA* **92**, 11259–11263 (1995).
- Hershko, A. & Ciechanover, A. The ubiquitin system for protein degradation. *Annu. Rev. Biochem.* **61**, 761–807 (1992).
- Alkalay, I. *et al.* *In vivo* stimulation of I κ B phosphorylation is not sufficient to activate NF- κ B. *Mol. Cell. Biol.* **15**, 1294–1301 (1995).
- DiDonato, J. A., Mercurio, F. & Karin, M. Phosphorylation of I κ B precedes but is not sufficient for its dissociation from NF- κ B. *Mol. Cell. Biol.* **15**, 1302–1311 (1995).
- Verma, I. M., Stevenson, J. K., Schwarz, E. M., Van Antwerp, D. & Miyamoto, S. Rel/NF- κ B/I κ B family: intimate tales of association and dissociation. *Genes Dev.* **9**, 2723–2735 (1995).
- Sun, S. C., Maggiorini, S. B. & Harhaj, E. Activation of NF- κ B by phosphatase inhibitors involves the phosphorylation of I κ B α at phosphatase 2A-sensitive sites. *J. Biol. Chem.* **270**, 18347–18351 (1995).
- Chen, Z. J., Parent, L. & Maniatis, T. Site-specific phosphorylation of I κ B α by a novel ubiquitination-dependent protein kinase activity. *Cell* **84**, 853–862 (1996).
- Haystead, C. M., Gregory, P., Sturgill, T. W. & Haystead, T. A. γ -Phosphate-linked ATP-Sepharose for the affinity purification of protein kinases. Rapid purification to homogeneity of skeletal muscle mitogen-activated protein kinase kinase. *Eur. J. Biochem.* **214**, 459–467 (1993).
- Connelly, M. A. & Marcu, K. B. CHUK, a new member of the helix-loop-helix and leucine zipper families of interacting proteins, contains a serine-threonine kinase catalytic domain. *Cell. Mol. Biol. Res.* **41**, 537–549 (1995).
- Baeuerle, P. A. & Henkel, T. Function and activation of NF- κ B in the immune system. *Annu. Rev. Immunol.* **12**, 141–179 (1994).
- Barnes, P. J. & Karin, M. Nuclear factor- κ B—A pivotal transcription factor in chronic inflammatory diseases. *New Engl. J. Med.* **336**, 1066–1071 (1997).
- Baeuerle, P. A. & Baltimore, D. NF- κ B: ten years after. *Cell* **87**, 13–20 (1996).
- Rothe, M., Wong, S. C., Henzel, W. J. & Goeddel, D. V. A novel family of putative signal transducers associated with the cytoplasmic domain of the 75 kDa tumor necrosis factor receptor. *Cell* **78**, 681–692 (1994).
- Hsu, H., Xiong, J. & Goeddel, D. V. The TNF receptor 1-associated protein TRADD signals cell death and NF- κ B activation. *Cell* **81**, 495–504 (1995).
- Hsu, H., Shu, B. H., Pan, M. G. & Goeddel, D. V. TRADD–TRAF2 and TRADD–FADD interactions define two distinct TNF receptor 1 signal transduction pathways. *Cell* **84**, 299–308 (1996).
- Cao, Z., Xiong, J., Takeuchi, M., Kurama, T. & Goeddel, D. V. TRAF6 in a signal transducer for interleukin-1. *Science* **283**, 443–446 (1996).
- Cao, Z., Henzel, W. J. & Gao, X. IRAK: a kinase associated with the interleukin-1 receptor. *Science* **271**, 1128–1131 (1996).
- Liu, Z.-G., Hu, H., Goeddel, D. V. & Karin, M. Dissection of TNF receptor 1 effector functions: JNK activation is not linked to apoptosis, while NF- κ B activation prevents cell death. *Cell* **87**, 565–576 (1996).
- Lee, F. S., Hagler, J., Chen, Z. J. & Maniatis, T. Activation of the I κ B α kinase complex by MEKK1, a kinase of the JNK pathway. *Cell* **88**, 213–222 (1997).
- Kallunki, T., Deng, T., Hibi, M. & Karin, M. c-Jun can recruit JNK to phosphorylate dimerization partners via specific docking interactions. *Cell* **87**, 929–939 (1996).
- Thompson, J. E., Philips, R. J., Erdjument-Bromage, H., Tempst, P. & Ghosh, S. I κ B β regulates the persistent response in a biphasic activation of NF- κ B. *Cell* **80**, 573–582 (1995).
- Kallunki, T. *et al.* JNK2 contains a specificity-determining region responsible for efficient c-Jun binding and phosphorylation. *Genes Dev.* **8**, 2996–3007 (1994).
- Helmberg, A., Auphan, N., Caelles, C. & Karin, M. Glucocorticoid-induced apoptosis of human leukemic cells is caused by the repressive function of the glucocorticoid receptor. *EMBO J.* **14**, 452–460 (1995).

Acknowledgements. We thank T. Thannhauser for help with peptide sequencing; C. Hoeger for the gift of γ -ATP–Sepharose; F. Mercurio and S. Ghosh for I κ B β plasmids; T. Haystead and L. Graves for discussion and for affinity resin; G. Walter for PP2A; E. Pasillas for oligonucleotide synthesis; G. Cadwell for technical assistance; and P. Alford for manuscript preparation. This work was supported by a Program Project grant from the National Cancer Institute and grants from the National Institutes of Environmental Health Sciences and the Human Frontier Science Project (HFSF). M.H., D.R. and E.Z. were supported by postdoctoral fellowships from HFSF, American Cancer Society (ACS) and ACS-California Division.

Correspondence and requests for materials should be addressed to M.K. (e-mail: karin_office@som-BSB.ucsd.edu).

A rotating disk of gas and dust around a young counterpart to β Pictoris

Vincent Mannings*, David W. Koerner† & Anneila I. Sargent*

* Division of Physics, Mathematics and Astronomy, California Institute of Technology, MS 105-24, Pasadena, California 91125, USA

† Jet Propulsion Laboratory, California Institute of Technology, MS 183-501, 4800 Oak Grove Boulevard, Pasadena, California 91109, USA

β Pictoris is the best known example of a main-sequence star encircled by a tenuous disk¹. Optical^{2,3} and infrared⁴ images of β Pic suggest that the disk is composed of dust grains which have been interpreted¹ as the debris generated by the disruption of the asteroid-sized remnants of planet-formation processes⁵. The star itself is relatively old, with an age in excess of 100 Myr. Here we present high-resolution millimetre-wave images of continuum and molecular-line emission from dust and gas surrounding a much younger star, MWC480: the stellar properties of MWC480 are similar to those of β Pic, but its age is just 6 Myr. The morphology of the circumstellar material and a comparison with the predictions of kinematic modelling indicate the presence of a rotating disk, gravitationally bound to the star. Moreover, the mass of the disk is greater than the minimum required to form a planetary system like our own⁵. We therefore suggest that the disk around the young star MWC480 could be a progenitor of debris disks of the type associated with older stars such as β Pic, and so holds much promise for the study of both the origin of debris disks and the early stages of the formation of planetary systems.

We observed MWC480 (HD31648) as part of a high-resolution millimetre-wave interferometric survey⁶ of young pre-main-sequence stars with spectral types Ae. These 'Herbig Ae' stars⁷ could be precursors to stars similar to β Pic, α Piscis Austrini (Fomalhaut), and α Lyrae (Vega), which also have early spectral types (A0–A5). A whole class of 'Vega-like'⁸ objects with spectral types B–M appear to be surrounded by grains distributed in low-mass disks^{1,9}. Optical and infrared images^{2–4} have directly demon-

strated the presence of a disk in the case of the A5 star β Pic, and analysis has suggested¹ that the grains comprising this and other Vega-like disks are the debris from continuing disruption of asteroid-sized planetesimals which are themselves perhaps the remnants of a previous epoch of planet formation. Our imaging survey aims to detect circumstellar disks associated with the much younger Herbig Ae stars, with a view to understanding the origins of debris disks like that surrounding β Pic.

MWC480 has a spectral type of A2e and is located in the Taurus-Auriga star-forming region at a distance of 140 pc. We determine a stellar mass $M_* = (2.3^{+0.1}_{-0.3})M_\odot$ (where $M_\odot$ is the solar mass), and model isochrones¹⁰ imply an age of ~ 6 Myr. Of our survey targets⁶, this source displayed the strongest continuum emission at wavelength $\lambda \approx 3$ mm and was moderately bright, but spatially unresolved, in $^{13}\text{CO}(1 \rightarrow 0)$ line emission. We followed our exploratory observations with higher-resolution measurements, which we present here. MWC480 was mapped with an angular resolution of $1.8'' \times 1.9''$ (full-width at half maximum, FWHM) in 1.3-mm continuum and $\text{CO}(2 \rightarrow 1)$ line emission on 15 October 1996 and 16 January 1997 using Caltech's Owens Valley Radio Observatory six-element millimetre-wave array at Big Pine, California. Visibility phases and amplitudes were calibrated using interleaved observations of the quasars 0528 + 134 and 3C111. The absolute flux density scale relied on measurements of Neptune. Respective uncertainties in fluxes and positions are $\sim 20\%$ and 0.5 arcsec.

An image of the 1.3-mm thermal continuum emission from dust grains around MWC480 is shown in Fig. 1a. The dust region is elongated along position angle (PA) 157^{+4}_{-3} degrees east of north, with an integrated flux density of 279 ± 7 mJy. Deconvolution from the beam assuming an elliptical gaussian source yields a major axis of 1.20 ± 0.15 arcsec (FWHM), corresponding to a half-maximum size $\sim 85 \pm 20$ AU.

Molecular-line emission from MWC480 is shown in Fig. 1b and appears more extended than the dust emission, with a brightness profile that is better resolved. Dust probably extends out to the same limits as the gas but, owing to different emissivity properties, it is not readily detectable at such distances from the star. The outer region of the molecular emission is at levels that are not well matched by a single gaussian component fit to the inner region. This peaked nature of the radial fall-off is as expected for a disk with radial density and temperature functions that vary as power laws. At the resolution attained in these observations, however, it can be

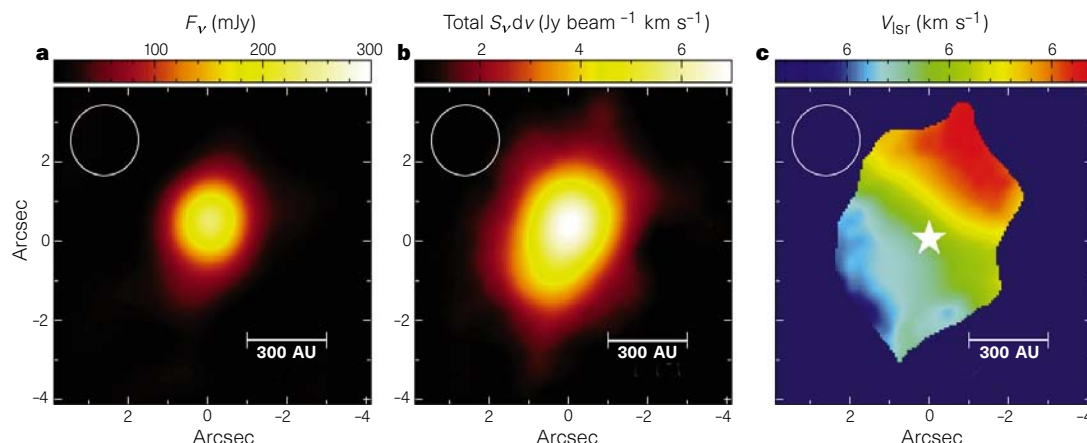


Figure 1 False-colour images of millimetre-wave emission from MWC480. In each panel, the full-width at half-maximum of the synthesized beam is shown at top left, and offsets in arcseconds from the stellar position are indicated along the vertical and horizontal axes. **a**, Thermal continuum from dust grains, at wavelength $\lambda = 1.3$ mm. The flux density (F_v) scale is shown at the top of the panel. **b**, $\text{CO}(2 \rightarrow 1)$ emission, integrated across the full velocity range ($+2.4$ to $+7.6$ km s^{-1})

in which emission is measured above the 3σ level. The scale at the top of the panel provides the integrated intensities ($S_v dv$). **c**, Intensity-weighted mean gas velocities at each spatial point of the structure shown in **b**. The stellar (systemic) velocity, $V_{\text{lsr}} \approx 5.1$ km s^{-1} , is represented by green. Blue-shifted (approaching) and red-shifted (receding) velocity components are shown as blue and red, respectively. The star symbol indicates the position of MWC480.

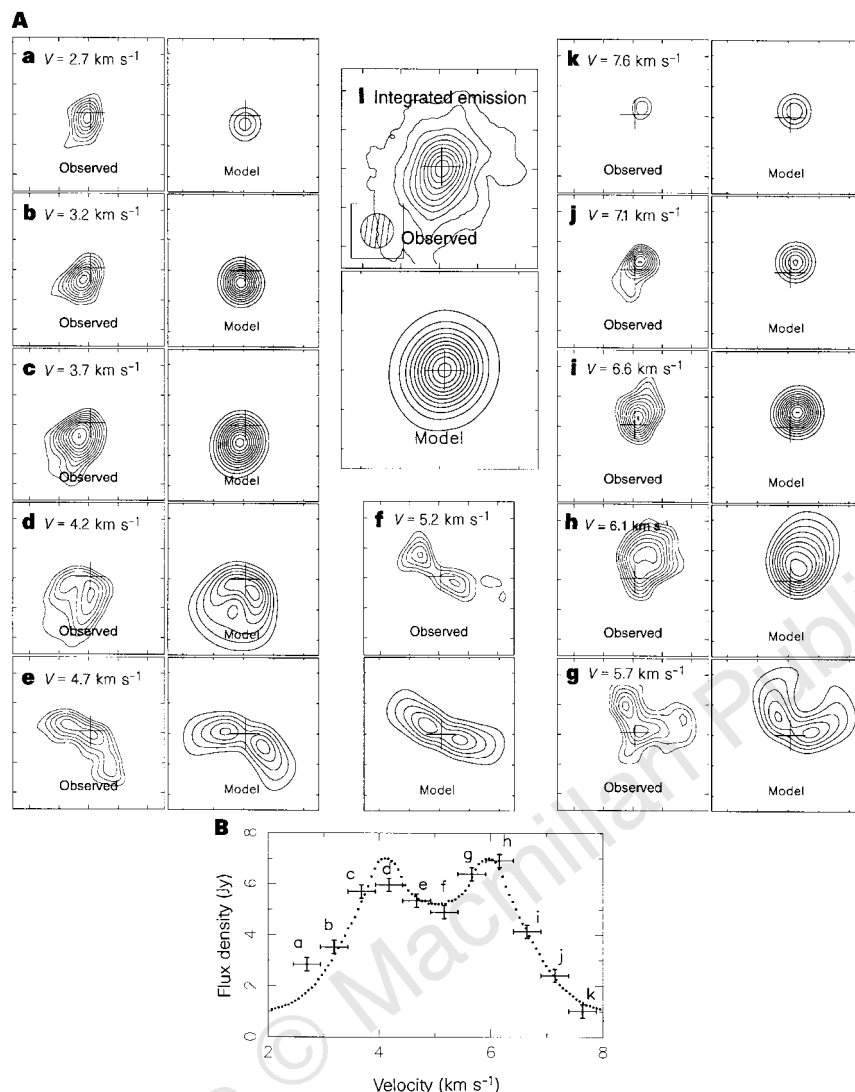


Figure 2 Aa–Al, Spectral-line maps of CO(2 → 1) emission from MWC480, shown adjacent to simulations of the emission predicted by a kinematic model of a disk in keplerian rotation with parameters as described in the text. Contour intervals are at the 1σ level, 170 mJy, and start at 3σ . Velocities labelled in the upper left-hand corner are referred to the local standard of rest (V_{lsr}). The best-fit model has a stellar (systemic) velocity corresponding to $V_{\text{lsr}} = 5.1 \text{ km s}^{-1}$. A contour plot of the velocity-integrated emission shown in Fig. 1b is shown in the upper central panel with contours spaced at $3\sigma = 0.63 \text{ Jy beam}^{-1} \text{ km s}^{-1}$; the full-width at half-maximum of the synthesized beam is shown in the lower left corner. Total emission from the best-fit model is contoured directly beneath. **B**, Plot of model spectrum together with the total flux from each of the maps **Aa–Ak**. Vertical error bars are derived from uncertainties in the pass-band calibration together with r.m.s. amplitude uncertainties in the maps. Horizontal error bars are the velocity width of emission in the spectral line maps.

well-approximated by the use of two gaussian components with best-fit half-maximum diameters of 1.9 and 6.3 arcsec, and peak amplitudes of 4.6 and $2.1 \text{ Jy beam}^{-1} \text{ km s}^{-1}$. The PA of both components is 157 ± 4 degrees and matches the orientation of the dust emission.

The line-of-sight velocity component of the circumstellar gas is revealed by the Doppler shift of the molecular line emission. Figure 1c shows the intensity-weighted mean gas velocity at each spatial point on the structure seen in Fig. 1b. Molecular gas southeast of the star seems to be approaching, whereas that to the northwest is receding, consistent with orbiting gas in a disk inclined with respect to the plane of the sky. However, to establish unambiguously the presence of a rotating disk it is necessary to analyse the velocity-dependent morphology and intensity of the gas emission in more detail.

Figure 2Aa–Ak displays the morphology of circumstellar emission in velocity intervals of width 0.49 km s^{-1} . For comparison, synthetic spectral-line maps, generated by a model of a disk in keplerian rotation¹¹, are shown in adjacent panels. In Fig. 2B, the locations of each velocity region are indicated on the spectrum of integrated intensities. Our model disk used a two-component gaussian emissivity distribution with half-maximum radii taken from the fit to the integrated map in Fig. 1b. The inclination angle i , outer cut-off radius R_D , and relative amplitudes of the gaussian emissivity components were varied in a least squares fit to the

spectral line maps. The central mass m (star + inner disk) was varied over a range consistent with estimates of the stellar mass reported above. Acceptable fits were found across the mass range $(2.0\text{--}2.4)M_\odot$, with corresponding inclinations $35\text{--}30^\circ$ and with ratios of the peak amplitudes of the inner and outer gaussian emissivity components as low as 8:1. For a central mass $m = 2.3M_\odot$, the best-fit parameter values are $i = 30^\circ$ and $R_D = 695 \text{ AU}$, with ratio of 9:1 for the inner-to-outer gaussian emissivity components. The resulting synthetic maps of a disk in keplerian rotation agree very well with the spectral-line maps of MWC480. Other interpretations of the gas motions, such as infall and outflow, are eliminated from consideration by this concordance¹¹.

The integrated CO line intensity, $23.2 \text{ Jy km s}^{-1}$, suggests¹² a mass of molecular hydrogen of $5.2 \times 10^{-5} M_\odot$, assuming low optical depths, an excitation temperature $T \approx 40 \text{ K}$, and a fractional CO abundance of $X(\text{CO}) = 10^{-4}$ relative to H_2 . Because the CO optical depths and the fractional abundance are very uncertain, this is, at best, a lower limit to the mass of the disk. A better mass estimate can be obtained from the 1.3-mm continuum flux density. We assume a standard grain opacity¹³, κ_ν , of $0.1(\nu/\text{GHz})/1,200^\beta \text{ cm}^2 \text{ g}^{-1}$, where ν is the observing frequency and β is the opacity index. We also adopt a gas-to-dust ratio of 100, by mass, and a value of unity^{14,15} for β . Then, for an optically thin source with a radially averaged temperature, T , of 40 K, we obtain a mass of gas and dust, $M \approx F_\nu d^2 / \kappa_\nu B_\nu(T) = 0.024 M_\odot$, where F_ν is the integrated continuum

flux density, d is the distance to the source, and B_ν is the blackbody Planck intensity. We can improve on this estimate by using a model^{14,15} in which radial disk temperature and density are allowed to vary as power laws. The exponent, 0.61, of the radial profile in disk temperature is obtained from a least-squares fit to the infrared spectral energy distribution; the radial density exponent¹⁵ is set at 7/4. Free parameters are disk mass, M_D , and β . From model fits to our 1.3- and 2.7-mm flux densities measured with the millimetre-wave array and to single-dish sub-millimetre flux densities (V. M. *et al.*, manuscript in preparation), we derive best-fit values for M_D and β of $(0.05 \pm 0.01)M_\odot$ and 1.0 ± 0.2 , respectively. We conclude that the true disk mass is a few per cent of a solar mass rather than the very low value derived from the CO line intensity. The opacity index, β , is significantly less than the value of 2 normally adopted¹⁶ for dust grains in the interstellar medium (ISM). This could be due to a population of grains in the MWC480 disk that are larger than 1 mm in size¹⁷, that is, several orders of magnitude greater than the sizes of ISM grains. Although this raises the possibility of grain accumulation, perhaps associated with the formation of larger bodies such as planetesimals, we note that opacity indices are also sensitive to chemical composition⁸ and to grain morphology¹⁹.

At $2.3M_\odot$, the MWC480 stellar mass is a factor of 3 greater than the typical value for T Tauri stars²⁰ (TTs). The TTs are young pre-main-sequence stars with later spectral types than the Herbig Ae stars (typically Me and Ke) and concomitantly lower masses, $\sim 0.7M_\odot$ on average²⁰. Millimetre-wave aperture synthesis images and kinematic models^{21,22} confirm that, as suggested by photometric observations²⁰, many TTs are accompanied by circumstellar disks that could be protoplanetary in nature^{23,24}. Our new observations of MWC480 provide the strongest evidence to date for a rotating disk around a higher-mass counterpart to the TTs.

The mass of the MWC480 disk, $(0.02\text{--}0.05)M_\odot$, is close to the peak of the mass distribution measured for disks of TTs²⁰. By contrast, disks around Vega-like stars are much less massive, by factors¹ of 10^{-5} to 10^{-6} . And, whereas the MWC480 and TT disks are dominated by gas, the Vega-like disks appear to be composed predominantly of grains²⁵. Indeed, the ages of Vega-like stars greatly exceed the timescales, in the absence of gas, for the removal of small inner-disk grains via radiation pressure and Poynting–Robertson drag. This has prompted suggestions^{1,26} that the grains are replenished by debris from continuing collisions and/or disruption of planetesimals. Could the MWC480 system represent the early stages of such a debris disk? With an age of 6 Myr, the massive rotating disk is a dense reservoir of orbiting material that could be sufficiently long-lived to support the growth of planetesimals. In fact the age of the disk exceeds, by two to three orders of magnitude, current estimates of the mean formation timescale^{5,27} for kilometre-sized planetesimals. Moreover, the mass of the MWC480 disk is greater than the minimum⁵ required to build a planetary system with an aggregate mass comparable to that of our Solar System. The disk encircling MWC480 could well be a progenitor of a debris disk such as that surrounding Pic, and provides an opportunity to examine the processes by which planetary systems are created. $\square$

Received 28 April; accepted 17 June 1997.

- Backman, D. E. & Paresce, F. in *Protostars & Planets III* (eds Levy, E. H. & Lunine, J.) 1253–1304 (Univ. Arizona Press, Tucson, 1993).
- Smith, B. A. & Terrile, R. J. A circumstellar disk around β Pictoris. *Science* **226**, 1421–1424 (1984).
- Golimowski, D. A., Durrance, S. T. & Clampin, M. Coronagraphic imaging of the β Pictoris circumstellar disk: evidence of changing disk structure within 100 AU. *Astrophys. J.* **411**, L41–L44 (1993).
- Lagage, P. O. & Pantin, E. Dust depletion in the inner disk of β Pictoris as a possible indicator of planets. *Nature* **369**, 628–630 (1994).
- Lissauer, J. J. Planet formation. *Annu. Rev. Astron. Astrophys.* **31**, 129–174 (1993).
- Mannings, V. & Sargent, A. I. A high-resolution study of gas and dust around young intermediate-mass stars: evidence for circumstellar disks in Herbig Ae systems. *Astrophys. J.* (in the press).
- Herbig, G. H. The spectra of Be- and Ae-type stars associated with nebulae. *Astrophys. J. Suppl. Ser.* **4**, 337–368 (1960).
- Aumann, H. H. *et al.* Discovery of a shell around α Lyr. *Astrophys. J.* **278**, L23–L27 (1984).
- Sylvester, R. J., Barlow, M. J., Skinner, C. J. & Mannings, V. Optical, infrared and millimetre-wave properties of Vega-like systems. *Mon. Not. R. Astron. Soc.* **279**, 915–939 (1996).
- D'Antona, F. & Mazzitelli, I. New pre-main-sequence tracks for $M \leq 2.5M_\odot$ as tests of opacities and convection model. *Astrophys. J. Suppl. Ser.* **90**, 467–500 (1994).

- Koerner, D. W. in *CO: Twenty-Five Years of Millimeter-Wave Spectroscopy* (eds Latter, W. B. *et al.*) 162–164 (IAU Symp. 170, Kluwer, Dordrecht, 1997).
- Scoville, N. Z. *et al.* High-resolution mapping of molecular outflows in GGC 2071, W49, and NGC 7538. *Astrophys. J.* **303**, 416–432 (1986).
- Hildebrand, R. H. Determination of cloud masses and dust characteristics from submillimetre thermal emission. *Q. J. R. Astron. Soc.* **24**, 267–282 (1983).
- Beckwith, S. V. W. & Sargent, A. I. Particle emissivity in circumstellar disks. *Astrophys. J.* **381**, 250–258 (1991).
- Mannings, V. & Emerson, J. P. Dust in disks around T Tauri stars: grain growth? *Mon. Not. R. Astron. Soc.* **267**, 361–378 (1994).
- Draine, B. T. & Lee, H. M. Optical properties of interstellar graphite and silicate grains. *Astrophys. J.* **285**, 89–108 (1984).
- Miyake, K. & Nakagawa, Y. Effects of particle size distribution on opacity curves of protoplanetary disks around T Tauri stars. *Icarus* **106**, 20–41 (1993).
- Pollack, J. B. *et al.* Composition and radiative properties of grains in molecular clouds and accretion disks. *Astrophys. J.* **421**, 615–639 (1994).
- Wright, E. L. Long-wavelength absorption by fractal dust grains. *Astrophys. J.* **320**, 818–824 (1987).
- Beckwith, S. V. W., Sargent, A. I., Chini, R. S. & Güsten, R. A survey for circumstellar disks around young stellar objects. *Astron. J.* **99**, 924–945 (1990).
- Koerner, D. W., Sargent, A. I. & Beckwith, S. V. W. A rotating gaseous disk around the T Tauri star GM Aurigae. *Icarus* **106**, 2–10 (1993).
- Dutrey, A., Guilloteau, S. & Simon, M. Images of the GG Tauri rotating ring. *Astron. Astrophys.* **286**, 149–159 (1994).
- Beckwith, S. V. W. & Sargent, A. I. Circumstellar disks and the search for neighbouring planetary systems. *Nature* **383**, 139–144 (1996).
- Sargent, A. I. in *Disks & Outflows around Young Stars* (eds Beckwith, S. V. W., Staude, J., Quetz, A. & Natta, A.) 1–23 (Springer, Berlin, 1996).
- Dent, W. R. F., Greaves, J. S., Mannings, V., Coulson, I. M. & Walther, D. M. A search for molecular gas components in prototypical Vega-excess systems. *Mon. Not. R. Astron. Soc.* **277**, L25–L29 (1995).
- Beust, H. & Lissauer, J. J. The effects of stellar rotation on the absorption spectra of comets orbiting β Pictoris. *Astron. Astrophys.* **282**, 804–810 (1994).
- Weidenschilling, S. The origin of comets in the solar nebula: a unified model. *Icarus* (in the press).

Acknowledgements. V.M. thanks M. Romans and N. Romans for their support. We thank the staff at the Owens Valley Radio Observatory for their assistance. The Owens Valley millimetre-wave array is supported by the NSF; array studies of young star and disk systems are supported in part by the Norris Planetary Origins project and by NASA's Origins of Solar Systems programme.

Correspondence should be addressed to V.M. (e-mail: vgm@astro.caltech.edu).

Observation of an internal wave attractor in a confined, stably stratified fluid

Leo R. M. Maas*, Dominique Benielli†, Joël Sommeria† & Frans-Peter A. Lam*

* Netherlands Institute for Sea Research, PO Box 59, 1790 AB Texel, The Netherlands

† Ecole Normale et Supérieure de Lyon, Laboratoire de Physique, 46 allée d'Italie, 69634 Lyon cedex 07, France

When a container of water is vibrated, its response can be described in terms of large-scale standing waves—the eigenmodes of the system. The belief that enclosed continuous media always possess eigenmodes is deeply rooted. Internal gravity waves in uniformly stratified fluids, however, present a counterexample. Such waves propagate at a fixed angle to the vertical that is determined solely by the forcing frequency, and a sloping side wall of the container will therefore act as a lens, resulting in ray convergence or divergence. An important consequence of this geometric focusing is the prediction¹ that, following multiple reflections, these waves will evolve onto specific paths—or attractors—whose locations are determined only by the frequency. Here we report the results of laboratory experiments that confirm that internal-wave attractors, rather than eigenmodes, determine the response of a confined, stably stratified fluid over a broad range of vibration frequencies. The existence of such attractors could be important for mixing processes in ocean basins and lakes, and may be useful for analysing oscillations of the Earth's liquid core and the stability of spinning, fluid-filled spacecraft.

The density gradient in a fluid is mainly determined by its temperature and dissolved salts. When the density increases at a constant rate in the direction of gravity the fluid, assumed to be

incompressible, is stably stratified and is characterized by a constant value of the stability frequency N . Regular perturbations of this static equilibrium, of frequency $\omega < N$, lead to internal gravity waves whose energy propagates obliquely² along straight lines—the rays, that make a fixed angle $\theta = \arccos(\omega/N)$ with the vertical. In our experiments, a uniformly stratified fluid is contained in a narrow tank, confining the wave propagation to a vertical plane. Tracing the rays that bounce back and forth between top and bottom boundaries and side walls is like following a frictionless ball on a billiard table, except that in this case the angle of reflection no longer needs to equal the angle of incidence as the ‘ball’ can only move in one of four directions, $\pm\theta \pmod{\pi}$, relative to the vertical. In a rectangular domain, whose boundaries are horizontal and vertical, there is no difference with the classical billiard table. Provided the depth-to-width ratio is rational, any ray travels along a closed path of finite length¹ and two rays retain their distance apart. In contrast, when one of the walls is sloping, two incident rays get focused or defocused on reflection from it (Fig. 1A). The case when these two processes offset each other, and all orbits are again closed (a ‘global resonance’), is exceptional. Also, chaotic rays, encountered in a variety of classical billiards³, are never obtained. Instead, in most cases focusing dominates and all ray paths become infinitely long, while approaching an attractor¹. An example is the (1,1)-attractor in Fig. 1A, the

square of heavy black lines, where an (m,n) -attractor is defined as having m reflections at the surface and n reflections at the vertical side-wall.

For a time-periodic disturbance, the spatial structure of the internal waves is a solution of a linear, hyperbolic (wave) equation in terms of the pressure and streamfunction fields², which, in turn, determine the velocity and density fields (Fig. 1B). This equation has solutions that are constant along lines, so-called characteristics, which coincide with the rays. Obtaining the ray paths—‘webs’ of connected characteristics—is therefore the first, geometric step in solving the hyperbolic equation. They span the fluid domain in a complete, but non-uniform way, as each of the rays approaches the attractor. Ray geometry shows that for any given frequency, the fluid domain is endowed with a structure (ultimately characterized by the shape of its attractor). The second step involved in solving this equation is contained in a description of the ‘dynamics’ on these webs, which visualizes (‘dresses’) this structure. Here, this amounts to the prescription of a ‘partial pressure’ f on each web, a quantity that is conserved along it. Pressure and streamfunction fields at an arbitrary point are then simply given by the sum and differences of the partial-pressure fields on the two characteristics that go through that point, and have a geometrical pattern that reproduces itself at smaller scales (Fig. 1B). The spatial structure is self-similar (fractal), and so is the dependence of the solution on dimensionless depth τ

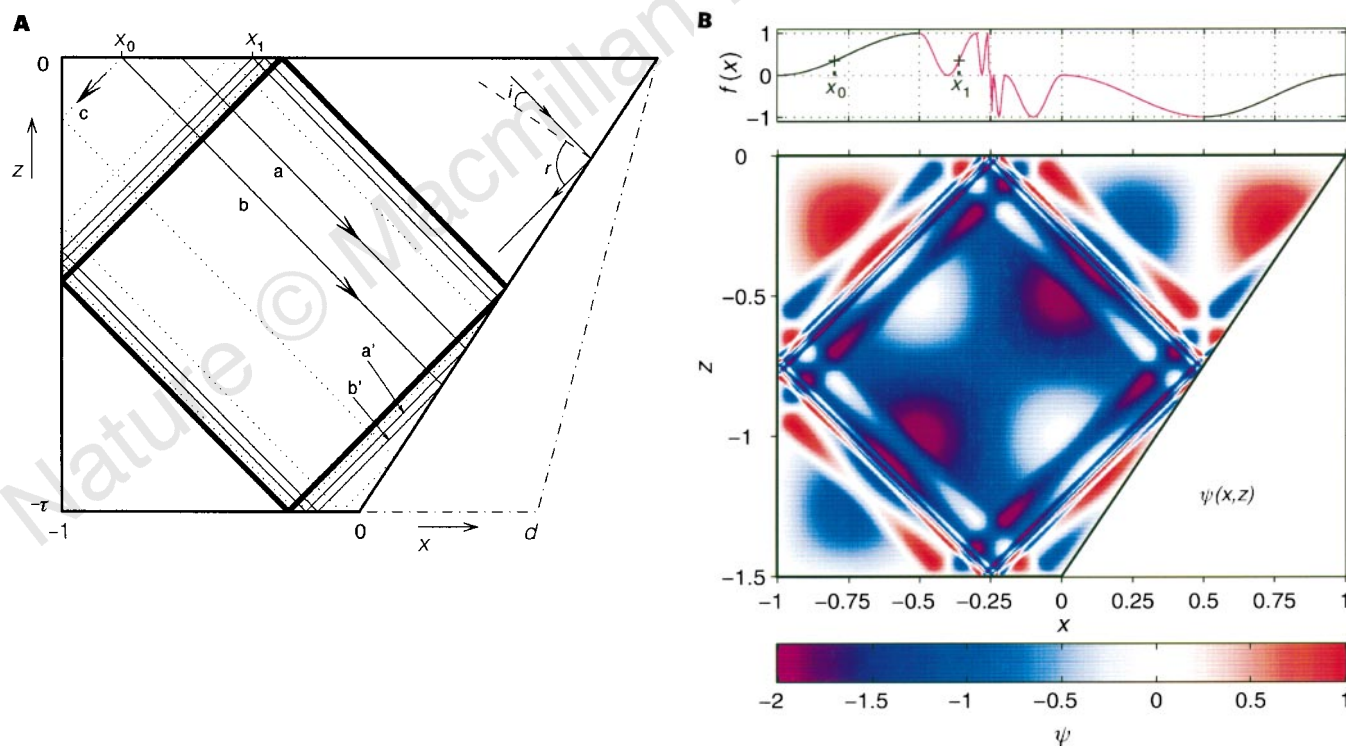


Figure 1 A, Sketch of basin with sloping side wall, varying in shape from triangular ($d = -1$) to rectangular ($d = +1$). Gravity acts vertically downwards. The general shape ($d \neq 0$) is dash-dotted, whereas the case $d = 0$, considered in the experiments, is given by solid lines. Coordinates x and z are scaled with half-width L and $L/|\tan \theta|$ respectively, so that true depth D is stretched to dimensionless depth $\tau = |\tan \theta| D/L = (N^2/\omega^2 - 1)^{1/2} D/L$ ($\tau = 3/2$ in this diagram). With this scaling, waves always propagate with an apparent inclination of 45° . Angle of reflection r on a sloping wall (with respect to its normal) therefore differs from angle of incidence i . When $r > i$, two parallel rays a and b , bouncing on a sloping wall, remain parallel after reflection, but get focused to a' and b' , eventually converging on the attractor, the square of heavy black lines. Defocusing is obtained, when $i > r$, as for rays a' and b' (reversing their direction), even though these rays eventually also approach the attractor, as ray c (dotted line), initially in a direction

opposite to b , shows. Surface intersections $x_0, x_1, \dots$ show convergence (to $x = -0.25$). **B**, The spatial part of the streamfunction $\psi(x, z) \exp(i\omega t)$, as governed by the hyperbolic equation $\psi_{xx} - \psi_{zz} = 0$, subject to the vanishing of ψ at the boundary (the impermeability condition). Here subscripts denote partial derivatives. A solution, vanishing at $z = 0$, is $\psi(x, z) = f(x - z) - f(x + z)$. The solution must also vanish at the other boundaries, which is satisfied when partial pressure f is invariant on each web of characteristics (lines $x \pm z = \text{constant}$)¹²¹. Then we can find $\psi(x, z)$ provided the value of $f(x)$ is prescribed on some appropriate intervals, as for the black line in the top panel within x intervals $[-1, -0.5]$ and $[0.5, 1]$. Starting, for instance, on the upper boundary at x_0 in the interval $[-1, -0.5]$, following connected characteristics, we return to x_1 in the neighbouring interval $[-0.5, -0.3]$, and infer $f(x_1) = f(x_0)$. By iterating this process for all points within these intervals we obtain f at each characteristic and hence the streamfunction.

(related to the wave period) and on the geometric parameter d (Fig. 2; these quantities are defined in Fig. 1).

The equation determining the spatial structure of the usual standing waves (for example, surface, or interfacial waves) is elliptic rather than hyperbolic and reduces to a standard eigenvalue problem. Solutions—eigenmodes, or seiches⁴—are then obtained only for a set of resonance frequencies, and they are smooth. These modes also dominate the response for arbitrary forcing frequencies. Here by contrast, we get solutions at any frequency below N and solutions are not smooth along the attractor (they are singular). For this reason the problem is often classified as ‘ill-posed’⁵. These solutions are still quite relevant physically, as shown by the laboratory experiments below.

A rectangular plexiglass container was filled with an exponentially stratified salt solution, such that the stability frequency was constant at $N = 1.89 \text{ s}^{-1}$. The container (shown in Fig. 3a) had a sloping side wall, extending from bottom centre to top right-hand corner yielding $d = 0$. Overall dimensions were: width (W) 96 mm; depth (D) 261 mm; and length ($2L$) 261 mm. Internal waves are visualized by the displacement of dye bands. These bands are produced by periodic dye injection into the brine entering the bottom of the tank, marking fluid parcels of given density. The deformation of the dye strips is visualized by the fluorescence excited by a vertical laser sheet.

The tank oscillates vertically with frequency 2ω (and amplitude Z), so as to produce a modulation of gravity (with amplitude $4Z\omega^2$). Any wave with frequency ω is then parametrically amplified, as for a pendulum. The waves are everywhere excited with the same phase, set by the forcing. The temporal evolution is therefore separable from the spatial structure, described by the hyperbolic equation. However, the amplitude is now growing exponentially in time, according to the canonical Mathieu equation, describing parametric excitation⁶.

Because of viscous effects, small-scale attractors are likely to be more suppressed, so our attempts were directed towards observing large-scale attractors. For the geometry described above, one expects to find a (1,1)-attractor in the interval of periods $1 \leq \tau \leq 2$, or, for the adopted aspect ratio $D/L = 2$, in the frequency interval $1/\sqrt{2} < \omega/N < 2/\sqrt{5}$.

We therefore explore this frequency range (keeping a constant excitation amplitude $Z = 10 \text{ cm}$). Approximately five minutes after the oscillation is started, a two-dimensional oscillatory motion of the dye lines with frequency ω becomes visible. This oscillation is localized, takes on a box-shaped form, and appears most pronounced around the predicted location of the (1,1)-attractor (Fig. 3a). In this initial growth phase, the oscillation has a standing nature: all points move up and down at the same time, while gradually growing in amplitude. Traces of strong shearing motion

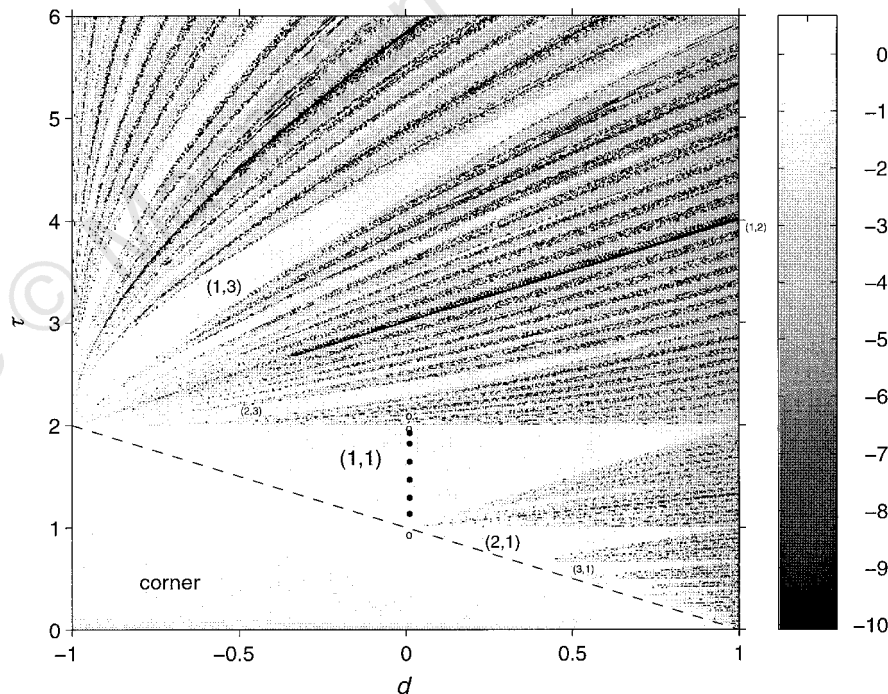


Figure 2 Convergence rate of neighbouring wave rays in parameter plane spanned by non-dimensional depth τ (related to internal wave period) and geometric parameter d (see definitions in Fig. 1A). As wave rays eventually almost always converge, the Lyapunov exponent λ , measuring this convergence, is generally negative and $\log_{10}(-\lambda)$ is therefore displayed. Light (dark) regions are strongly (weakly) convergent. Regions and lines are labelled with integer pairs (m, n) , which describe the number of reflections at the surface (m) and vertical side wall (n) of the attractor, or global resonance. For an (m, n) -attractor only one ray (the attractor) is truly periodic, for a global resonance every ray is. Continuous, light-coloured regions, identifying large-scale attractors (low m and n), are separated by darker regions where fine-scale attractors reign. Within these ranges still finer-scale attractors reside, ultimately approaching the black lines for

which there is no convergence at all (zero Lyapunov exponent): the set of global resonances. Wave rays may also be attracted to the right-hand upper corner of the basin in Fig. 1A, but this occurs only when the bottom slope is less steep than the rays (a subcritical slope)²², below the dashed line $\tau = 1 - d$. Above the critical curve, (m, n) -attractors can be identified. Remarkably, only attractors with an odd number of ‘vertical’ cells n , appear. The ‘missing’, even-numbered cells, correspond to the (m, n) -global resonances. The first of these, the (1, 2)-global resonance, is clearly identified by the black line $\tau = 3 + d$. Note that (m, n) -attractors, as well as (m, n) -global resonances, approach the classical (m, n) -cellular solutions of the form $\psi = \sin[m\pi(x + 1/2)]\sin[n\pi z/\tau]$ of the rectangle, which are found at $\tau = 2n/m$, as $d \rightarrow 1$. Filled and open circles at $d = 0$ indicate laboratory experiments in which ‘focusing modes’ were, or were not observed.

can be detected by the occurrence of sharp gradients in dye concentrations, especially near the part of the attractor closest to the sloping side wall (where the focusing takes place). Dye mixing is later observed in this zone.

This structure contrasts strongly with the resonance modes previously observed with the same apparatus⁷ in a rectangular geometry (Fig. 3b). Moreover, the wave attractor changes smoothly with frequency, whereas the resonance mode was obtained only in a narrow frequency range (depending on the forcing amplitude, as an instability 'tongue'). Instead, the attractor is observed in the whole predicted frequency interval, as indicated by filled circles in Fig. 2. Outside this interval (open circles), we observe no wave excitation, or complex (unidentified) three-dimensional modes.

Visualization of the attractor can be improved by subtracting the initial state (where dye lines are flat), as in Fig. 4a for the standing

phase. After about two minutes, the standing wave gives way to the permanent regime, in which the amplitude of the internal wave saturates. The wave then takes on a propagating character (Fig. 4b), while it displays some chaotic particle motion. Phase propagation is perpendicular to the attractor, as clearly observed by tracking a wave nodal line (having zero elevation); this is consistent with an energy flux directed parallel to the attractor, in the clockwise (focusing) direction.

The concurrently observed irregular behaviour near the slope could be related to chaotic particle transport⁸ by the propagating wave. The standing-wave regime by contrast is very regular. Particle motion is then periodic, the streamfunction acting like the hamiltonian of a system with one degree of freedom.

A stratified fluid in a non-trivially shaped container possesses inviscid modes of oscillation which are singular along a wave

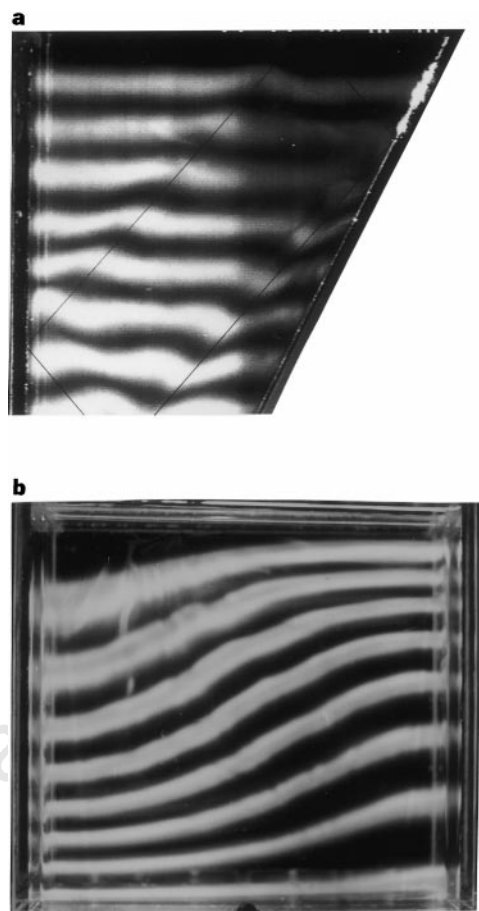


Figure 3 a, Side view of laboratory tank, showing a maximum displacement of (initially horizontal) dye lines during the growth phase (standing wave), 9 minutes after the oscillation of the table was started. As expected, the displacement is strongest along the attractor, represented by the solid line (tank oscillation amplitude $Z = 10$ cm, $\tau = 1.71$). Note that the bottom of the tank is located slightly below the lower edge of the figure. The occurrence of wave growth depends on the frequency and amplitude of excitation and is apparently inhibited by viscous effects when the attractor shape becomes either too complicated, or too narrow, as when the box-shaped attractor degenerates into a line. Thus for this tank ($d = 0$), in the frequency interval for which the (1, 1)-attractor is obtained, the instability threshold is close to $Z \approx 15$ cm near the bounds of the interval, but is lower, $Z \approx 8$ cm at its middle. **b**, As **a**, but with a square domain (from ref. 7), for which an eigenmode structure is obtained, in sharp contrast with the attractor observed in the present study.

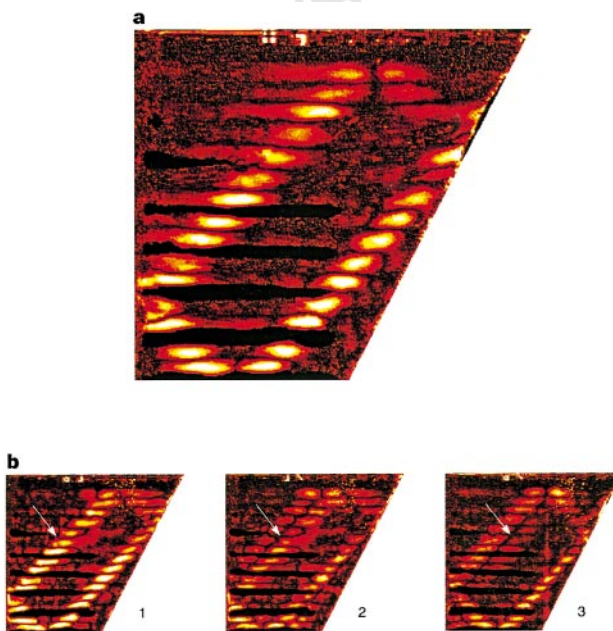


Figure 4 a, Side view of tank showing the difference between (maximally) displaced dye lines and their horizontal initial state (using computer processing) in the standing phase, 9 minutes after oscillation of the tank was started ($Z = 10$ cm, $\tau = 1.74$). **b**, As **a**, but in the permanent regime (10 minutes after the start of oscillation) showing a propagating wave. Energy propagation of internal gravity waves is perpendicular to phase propagation (such that their vertical components are opposite). Thus, in Fig. 1A, when phase propagates from b to a, energy propagates along these phase-lines, in the direction of the arrows. In the permanent regime, phase propagation perpendicular to the attractor is observed on each of the four sides of the box-shaped attractor in a manner consistent with clockwise energy propagation around it. This is seen by tracking (black) nodal lines on successive views (indicated by arrows on one side of the box). Integer labels n (where $n = 1, 2$ and 3) refer to relative time $T \times n/12$, where $T = 4.37$ s is the wave period. Propagating waves are theoretically obtained for complex f . For each solution, there is also a complex conjugate solution, representing a wave propagating in the opposite direction. Each wave has an infinite path, and can be likened to waves on an infinite string. But both waves approach the same attractor, as illustrated by rays b and c in Fig. 1A. In the growth phase these oppositely propagating waves combine to form a standing wave, whose uniform phase allows effective growth by parametric instability. Once these waves reach the vicinity of the attractor they transfer energy to smaller scales where it is selectively withdrawn by nonlinear and viscous processes. This provides the additional damping required for saturation of the parametric instability and breaks the symmetric propagation of energy through the 'string', thus giving the final wave pattern its propagating character.

attractor. Our experiments confirm the physical relevance of this attractor, in spite of viscous and nonlinear effects. A standing wave is initially excited by the parametric instability, turning into a propagating wave in the permanent regime. A usual boundary forcing should only produce the latter regime, with permanent excitation at large scale, and energy transfers towards small scales along the attractor⁹. The system also has a broad-band response in frequency, set by the interval of existence of the attractor. This behaviour contrasts dramatically with the narrow peaks of usual resonances, whose width is set by dissipative effects.

Similar internal wave attractors could be predicted, by ray tracing, in ocean basins or lakes. Waves must be amplified by focusing at their contact with the sloping bottom. This should determine zones of intense mixing, favouring life, by providing a pathway for the transport of nutrients and rare metals from the bottom to the photic zone^{10,11}. This focusing process has often been invoked^{12–14}, but without considering its localization on a global attractor. Such localization depends of course on the wave frequency. In the ocean, important contributions at the tidal frequency should be expected, but, unlike our experiments, wind-related frequencies are simultaneously excited, and the widest frequency 'windows', corresponding to an attractor related to the particular geometry, may show up. The dressing of the attractor by the wave solution is clearly understood in our two-dimensional case, but a (non-trivial) three-dimensional generalization must be sought for natural basins.

Our results could also be relevant to an experiment¹⁵ that showed inertial oscillations in a fluid contained in a truncated cone rotating around its axis (where axisymmetric modes are excited by a sinusoidal modulation of the rotation rate). The spectral response was in this case interpreted in terms of resonance peaks of standing waves (presumably broadened by viscous effects). We note, however, that the spectral 'peaks' and 'valleys' coincide remarkably well with the regions of strong and weak convergence rates (obtained as an appropriate cross-section in Fig. 2 for $d = -0.478$): so we are led to conclude that the broadness of these response 'peaks' may well reflect a genuine, finite bandwidth of the response rather than a viscously spread-out resonance spike.

The problem of confined inviscid inertial oscillations, as occur in a spherical shell, is also ill-posed¹⁶ (unlike in a sphere). Smooth eigenmodes can be computed by introducing a small viscosity¹⁷, but better physical insight is expected in terms of wave attractors. The simplest attractors would again show up as wide resonance bands, whose width is independent of dissipative effects. This may be relevant to analysing inertial oscillations of the Earth core^{18,19} and to the stability of liquid-filled, rotating spacecraft²⁰. □

Received 19 December 1996; accepted 17 June 1997.

1. Maas, L. R. M. & Lam, F.-P. A. Geometric focusing of internal waves. *J. Fluid Mech.* **300**, 1–41 (1995).
2. Görtler, H. Über eine Schwingungserscheinung in Flüssigkeiten mit stabiler Dichteschichtung. *Z. Angew. Math. Mech.* **23**, 65–71 (1943).
3. Berry, M. V. Regularity and chaos in classical mechanics, illustrated by three deformations of a circular 'billiard'. *Eur. J. Phys.* **2**, 91–102 (1981).
4. Wedderburn, E. M. The temperature seiche. I–IV. *Phil. Trans. R. Soc. Edinb.* **47**, 619–642 (1911).
5. Courant, R. & Hilbert, D. *Methoden der Mathematischen Physik* Vol. II (Springer, Berlin, 1937).
6. McEwan, A. D. & Robinson, R. M. Parametric instability of internal gravity waves. *J. Fluid Mech.* **67**, 667–687 (1975).
7. Benielli, D. & Sommeria, J. Excitation of internal waves and stratified turbulence by parametric instability. *Dyn. Atmos. Oceans* **23**, 335–343 (1996).
8. Joseph, B. Chaotic mixing by internal inertia-gravity waves. *Phys. Fluids* **9**, 945–962 (1997).
9. Phillips, O. M. Energy transfer in rotating fluids by reflection of inertial waves. *Phys. Fluids* **6**, 513–520 (1963).
10. Steele, J. H. *Spatial Pattern in Plankton Communities* (Plenum, New York, 1978).
11. Holligan, P. M., Pingree, R. D. & Mardell, G. T. Oceanic solitons, nutrient pulses and phytoplankton growth. *Nature* **314**, 348–350 (1985).
12. Eriksen, C. Implications of ocean bottom reflections for internal wave spectra and mixing. *J. Phys. Oceanogr.* **15**, 1145–1156 (1985).
13. Garrett, C. J. R. & Gilbert, D. In *Small-scale Turbulence and Mixing in the Ocean* (eds Nihoul, J. C. J. & Jamart, B. M.) 405–423 (Elsevier, Amsterdam, 1988).
14. Ivey, G. N. & Nokes, R. I. Vertical mixing due to the breaking of critical internal waves on sloping boundaries. *J. Fluid Mech.* **204**, 479–500 (1989).
15. Beardsley, R. C. An experimental study of inertial waves in a closed cone. *Stud. Appl. Math.* **49**, 187–196 (1970).
16. Stewartson, K. & Rickard, J. A. Pathological oscillations of a rotating fluid. *J. Fluid Mech.* **35**, 759–773 (1969).

17. Ricourt, M. Inertial modes in the liquid core of the earth. *Phys. Earth Planet. Inter.* **91**, 41–46 (1995).
18. Melchior, P. & Ducarme, B. Detection of inertial gravity oscillations in the Earth's core with a superconducting gravimeter at Brussels. *Phys. Earth Planet. Inter.* **42**, 129–134 (1986).
19. Aldridge, K. D. & Lumb, L. I. Inertial waves identified in the Earth's fluid outer core. *Nature* **325**, 421–423 (1987).
20. Kobine, J. J. Inertial wave dynamics in a rotating and precessing cylinder. *J. Fluid Mech.* **303**, 233–252 (1995).
21. Magaard, L. Ein Beitrag zur Theorie der internen Wellen als Störungen geostrophischer Strömungen. *Deutsche Hydrogr. Z.* **21**, 241–278 (1968).
22. Cacchione, D. & Wunsch, C. Experimental study of internal waves over a slope. *J. Fluid Mech.* **66**, 227–239 (1974).

Acknowledgements. We thank G. van der Plas and S. Dalziel for help with post-processing of the data, and C. Staquet and B. Voisin for suggestions.

Correspondence should be addressed to L.R.M.M. (e-mail: maas@nioz.nl).

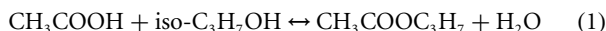
Discovery of a reactive azeotrope

W. Song, R. S. Huss, M. F. Doherty & M. F. Malone

Department of Chemical Engineering, Goessmann Laboratory,
University of Massachusetts, Amherst, Massachusetts 01003–3110, USA

Mixtures are azeotropic if they can be distilled (or condensed) without a change of composition¹. The existence of azeotropes in multicomponent mixtures in the absence of chemical reactions is well understood phenomenologically^{2,3} and theoretically^{4,5}. Azeotropes place a fundamental limit on the compositions attainable in mixtures by fractional distillation, but they can in some cases be 'broken' by carrying out chemical reaction and separation simultaneously rather than sequentially^{6–9}. Here we report the discovery of a boiling state of constant composition and temperature in a mixture of acetic acid, isopropanol, isopropyl acetate and water that is simultaneously in both reaction and phase equilibrium. These states, which we call reactive azeotropes, were predicted recently^{10,11}. Without reaction, the mixture exhibits three two-component azeotropes, one three-component azeotrope but no four-component azeotrope; the last appears only under equilibrium reaction conditions. These findings may constrain technologies in which reaction and separation are conducted simultaneously, for example by limiting the conditions under which an azeotrope can be broken by chemical reactions to yield a high-purity product. In other cases the presence of a reactive azeotrope may be advantageous⁹.

Without chemical reaction, a mixture of acetic acid, isopropanol, isopropyl acetate and water has three binary azeotropes and a ternary azeotrope as listed in Table 1. There is no four-component azeotrope in this mixture. An acid-catalysed esterification reaction occurs in this mixture:



This reaction occurs slowly; the rate can be increased substantially by the addition of a strong acid catalyst. The non-reactive binary azeotropes between water and isopropanol and between isopropyl acetate and isopropanol will also exist in the reacting mixture. However, a binary mixture of isopropyl acetate and water, as well as any mixture containing three of the components, will react and the corresponding azeotropes are absent.

A recent and general theory for mixtures in simultaneous vapour–liquid and chemical reaction equilibrium revealed the possibility of a state with a constant boiling temperature and constant, though not equal, vapour and liquid compositions^{10,11}. In a reacting, boiling liquid each component has a rate of vaporization and a rate of creation or loss due to reaction. When the net rate for each component is zero, an azeotrope occurs. However, the existence of a reactive azeotrope has not been confirmed experimentally.

The theory^{10,11} provides a general transformation of variables

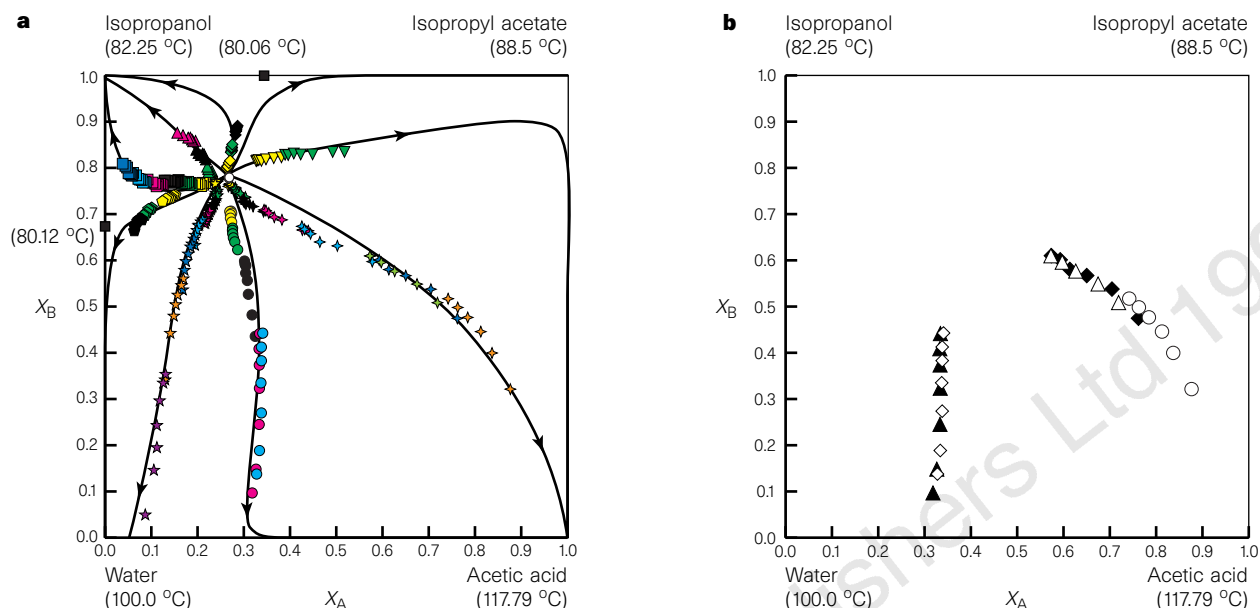


Figure 1 a, Comparison of experiments and calculated residue curves; X_A and X_B are composition variables (see text). The calculated residue curves are shown by the solid lines. The arrows show the direction of increasing time and temperature. The shapes of the plotting symbols show experiments within a series; the colours of the symbols distinguish different runs within that series. The run closest to the azeotrope is shown with open circles. The temperatures shown are those predicted by the model. The calculated curves were generated with the Non-Random Two Liquid (NRTL) model¹⁴ for activities in the liquid phase, including association of acetic acid in the vapour phase. Parameters for the model were taken from the literature¹² except for the isopropanol-isopropyl acetate and water-isopropyl acetate interactions, which we estimated from the UNIFAC group contribution method¹⁵ optimized for the locations of the non-reactive

azeotropes listed in Table 1. We used a constant value of 8.7 for the equilibrium constant for the reaction (see equation (1)) as reported by Lee and Kuo¹³. **b**, Replicated experiments. The filled diamonds and triangles show the two original experiments and the open diamonds and triangles show the replications. The open circles are a run beginning from a composition near the last measured value of the original series. In the original and replicated experiments, measurements were taken at the same time intervals. We note that the replicated runs might be slightly shifted in time owing to small differences in vapour removal rate. However, as demonstrated in the figure, the trajectories of compositions correspond quite closely to those expected for systems with dynamics that follow autonomous models such as equation (3).

from mole fractions (x) to a new set of composition variables (X). This takes into account the reduction in degrees of freedom due to chemical reaction and provides a natural coordinate system. In the mixture studied here, the transformed variables represent the compositions in terms of the total amount of acetic acid and isopropanol, including the equivalent amounts present in the other two components. These are:

$$\begin{aligned} X_A &= x_{\text{acetic acid}} + x_{\text{i-propyl acetate}} \\ X_B &= x_{\text{i-propanol}} + x_{\text{i-propyl acetate}} \\ X_C &= x_{\text{water}} - x_{\text{i-propyl acetate}} \end{aligned} \quad (2)$$

The feasible composition space is the square shown in Fig. 1a. Each vertex of the square is a pure component, each edge is a non-reactive binary mixture, and any point inside the square is a four-component reactive mixture.

Figure 1a also shows residue curves, which represent the liquid composition paths in an isobaric open evaporation. For a single chemical reaction, these are the solutions of

$$\frac{dX_i}{d\xi} = Y_i - X_i \quad (3)$$

for $i = 1, 2, \dots, c - r - 1$ where c is the number of chemical species, r is the number of independent chemical reactions, and ξ is a dimensionless time. The transformed liquid and vapour compositions, X and Y , are related by phase and chemical-reaction equilibrium.

Residue curves begin and end at azeotropes or pure components and are directed from low to high temperatures. In Fig. 1a, all of the calculated residue curves (solid lines) originate from a point inside the figure where

$$X_i = Y_i \quad (4)$$

for $i = 1, 2, \dots, c - r - 1$. This point is a minimum-boiling, reactive azeotrope with a molar liquid composition of 5.30% acetic acid, 56.55% isopropanol, 21.87% isopropyl acetate and 16.28% water at a temperature of 78.99 °C and a vapour composition of 0.33% acetic acid, 51.58% isopropanol, 26.84% isopropyl acetate and 21.25% water. The corresponding transformed compositions are $X_A = Y_A = 0.2717$ and $X_B = Y_B = 0.7842$.

To test this prediction, we measured the composition changes in a liquid undergoing simultaneous open evaporation and chemical reaction. A solid acidic resin (Amberlyst 15, Rohm & Haas Co., Philadelphia, PA) was used as a catalyst. The liquid compositions were measured periodically by removing small samples for analysis by gas chromatography (Hewlett-Packard 5890 Series II gas chromatograph, TC detector, HP-INNOwax column). For gas chromatography measurements, 0.1 ml of liquid sample was diluted with 20 ml of anhydrous diethylether. A small quantity (1 μ l) of the diluted solution was injected for analysis. The response of the gas chromatograph was calibrated for each of the four components of the mixture. At the beginning of an experiment, the mixture was refluxed with catalyst until the temperatures of both the liquid and vapour phases were stabilized. Vapour was then removed continuously and the liquid compositions were measured as a function of

Table 1 Azeotropes at 1 atm

Boiling temperature (°C)	Mole fractions			
	Acetic acid	Isopropanol	Isopropyl acetate	Water
75.5	0.00	0.1377	0.4737	0.3887
76.6	0.00	0.00	0.5980	0.4020
80.1	0.00	0.6480	0.3520	0.00
80.1	0.00	0.6873	0.00	0.3127

Data reported by Horsley³.

time. Eventually, the amount of liquid residue was small enough to preclude sampling, ending the experiment. Subsequent experiments, beginning at compositions close to the end point of the previous experiment, were necessary in most cases to get a clear measure of each residue curve.

We studied the effect of catalyst concentration by using the same initial composition of the mixture with different amounts of catalyst. Our interest here is in the limiting case approaching phase and reaction equilibrium so we sought to find a large amount of catalyst, a low rate of vapour withdrawal, and the appropriate start-up procedure to approach this condition. With 27.5 g of catalyst (0.135 mol H⁺) in an initial volume of 306 ml liquid, the theoretical residue curve is approached closely. The similarity between the experimental residue curves for 22.0 and 27.5 g of catalyst showed an approach to this limiting case. Based on these results, we used 30.0 g catalyst (0.147 mol H⁺) in an initial volume of 306 ml for all experiments reported here. Using a similar approach, we determined a heating rate low enough to ensure that the results are independent of the rate of mass transfer between liquid and vapour phases. In our experiment, this was achieved in a time of ~6 hours to remove ~75% of the initial liquid or an average vapour removal rate of ~38 ml h⁻¹.

A series of nine experiments with initial conditions in the neighbourhood of the calculated azeotrope were performed in a total of 39 runs; each set in a series was intended to represent one residue curve; successive runs were necessary in most cases to show a significant portion of each curve. The results are shown in Fig. 1a. The open circles indicate the position of the reactive azeotrope as closely as it could be determined; these points are nearly overlapping as predicted. This single run consists of five data points taken over a span ~4 hours, during which time the compositions remained nearly constant and the boiling temperature varied by <0.1 °C. The measured molar liquid composition of the reactive azeotrope was 5.4% acetic acid, 56.5% isopropanol, 21.4% isopropyl acetate and 16.7% water at a temperature of 78.6 °C, with a corresponding vapour composition of 0.0% acetic acid, 49.1% isopropanol, 27.0% isopropyl acetate and 23.9% water. The corresponding transformed compositions are $X_A = 0.268$, $X_B = 0.779$ and $Y_A = 0.270$, $Y_B = 0.761$.

We assessed the reproducibility of the results by repeating two runs and found the data to be highly reproducible, as shown in Fig. 1b.

These results demonstrate the existence of a reactive azeotrope not present in the non-reactive mixture. This confirms the predictions of a model based on classical thermodynamics and provides a basis for predicting the behaviour of other mixtures. For example, we also predict reactive azeotropes in the esterification of acetic acid with *n*-propanol or iso-butanol, as well as a near reactive azeotrope in the reaction of methanol and isobutene to form methyl *t*-butyl ether (MTBE). In the production of MTBE the 'reactive azeotrope' is predicted to contain a substantial amount of MTBE and is responsible for making reactive distillation the preferred process technology⁹. We believe that reactive azeotropes are common and for some systems can determine the optimal choice of process technology. As a further example, the manufacture of methyl acetate uses reactive distillation, because reaction eliminates the azeotrope

between water and methyl acetate that otherwise limits the purity of the product^{8,9}. But our results show that a similar process will not work for isopropyl acetate because of the existence of a reactive azeotrope in the four-component reactant/product mixture. □

Received 20 January; accepted 24 June 1997.

- Rowlinson, J. S. *Liquids and Liquid Mixtures* 2nd edn (Butterworth, London, 1969).
- Lecat, M. *Tables Azeotropiques: I. Azeotropes Binaires Orthobares* (l'Auteur, Bruxelles, 1949).
- Horsley, L. H. *Azeotropic Data-III* (Am. Chem. Soc., Washington DC, 1973).
- Malesinski, W. *Azeotropy and Other Theoretical Problems of Vapour-Liquid Equilibrium* (Interscience, New York, 1965).
- Swietoslawski, W. *Azeotropy and Polyazeotropy* (Pergamon, Warsaw, 1963).
- Xu, X., Zhu, B., Chen, H. & Bai, G. Reactive distillation (I). *Shiyou Huagong (Petrochemical Technology)* **14**, 480–486 (1985).
- Xu, X., Zhu, B., Chen, H. & Bai, G. Reactive distillation (II). *Shiyou Huagong (Petrochemical Technology)* **14**, 550–555 (1985).
- Agreda, V. H. & Partin, L. R. Reactive distillation process for the production of methyl acetate. US Patent No. 4435595 (to Eastman Kodak Company) (1984).
- Doherty, M. F. & Buzad, G. Reactive distillation by design. *Trans. Inst. Chem. Eng. A* **70**, 448–458 (1992).
- Barbosa, D. & Doherty, M. F. Theory of phase diagrams and azeotropic conditions for two-phase reactive systems. *Proc. R. Soc. Lond. A* **413**, 443–458 (1987).
- Ung, S. & Doherty, M. F. Necessary and sufficient conditions for reactive azeotropes in multireaction mixtures. *Am. Inst. Chem. Eng. J.* **41**, 2383–2392 (1995).
- Gmehling, J. & Onken, U. *Vapor-Liquid Equilibrium Data Collection* (DECHEMA Chemistry Data Ser., Frankfurt, 1977).
- Lee, L.-S. & Kuo, M.-Z. Phase and reaction equilibrium of the acetic acid-isopropanol-isopropyl acetate-water system at 760 mm Hg. *Fluid Phase Equil.* **123**, 147–165 (1996).
- Renon, H. & Prausnitz, J. M. Local compositions in thermodynamic excess functions for liquid mixtures. *Am. Inst. Chem. Eng. J.* **14**, 135–144 (1968).
- Fredenslund, A., Gmehling, J. & Rasmussen, P. *Vapor-Liquid Equilibria Using UNIFAC* (Elsevier, Amsterdam, 1977).

Acknowledgements. This work was supported by National Science Foundation and the National Environmental Technology Institute.

Correspondence should be addressed to M.F.D. (e-mail: mdoherty@ecs.umass.edu).

Decadal predictability of North Atlantic sea surface temperature and climate

R. T. Sutton*† & M. R. Allen*‡

* *Atmospheric Oceanic and Planetary Physics, Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, UK*

† *Centre for Global Atmospheric Modelling, Department of Meteorology, University of Reading, PO Box 243, Earley Gate, Reading RG6 6BB, UK*

‡ *Space Science Department, Rutherford Appleton Laboratory, Chilton OX11 0QX, UK*

The weather at middle latitudes is largely unpredictable more than a week or so in advance, whereas fluctuations in the ocean may be predictable over much longer timescales. If decadal fluctuations in North Atlantic sea surface temperature^{1–6} could be predicted, it might be possible to exploit their influence on the atmosphere^{7–10} to forecast decadal fluctuations in climate¹¹. Here we report analyses of shipboard observations that indicate significant decadal predictability of North Atlantic sea surface temperature, arising from the advective propagation of sea-surface-temperature anomalies⁴ and the existence of a regular period of 12–14 years in the propagating signals. The same timescale can be identified in a dipole-like pattern of North Atlantic sea-level pressure variability^{1,7,12}. We propose a mechanism which may connect these oceanic and atmospheric fluctuations, possibly as part of a coupled ocean–atmosphere mode of variability⁷. Our results are encouraging for the prospects of forecasting natural fluctuations in the climate of the North Atlantic region several years in advance.

Several authors have suggested that much of the decadal variability in North Atlantic sea surface temperature (SST) can be explained as a local oceanic response to atmospheric variability^{1,3,13},

whereas others have argued that non-local dynamical processes in the ocean must play an important role^{2,4,14}. In the latter group, Hansen and Bezdek⁴ have recently presented evidence of advection of SST anomalies, especially along the path of the North Atlantic Current (NAC), an extension of the Gulf Stream. The importance of this suggestion is that fluctuations in SST that arise from advection may be predictable. We use analyses of monthly mean SST compiled from ship observations for the period 1945–89¹⁵ to investigate the evidence for predictability in North Atlantic SST associated with oceanic advection, and examine whether advection could be an important element in a possible coupled ocean–atmosphere mode of North Atlantic variability⁷.

Figure 1a shows an estimate of the upper-ocean circulation in the North Atlantic¹⁶. Figure 1b shows the correlation between low-frequency fluctuations in local wintertime SST and low-frequency fluctuations in wintertime SST averaged over the region 80–60° W, 31.5–38.5° N (the vicinity of Cape Hatteras) as a function of lag in years. SST anomalies are most likely to express deeper subsurface

anomalies in winter, when the oceanic mixed layer is at its deepest. Coherent SST fluctuations appear to propagate from the coast of North America across the Atlantic to the northwest of Scotland, following the strong SST gradients that are associated with the Gulf Stream and NAC.

In Fig. 2 we show the propagating SST anomalies that give rise to the correlations seen in Fig. 1b. A series of warm and cold anomalies propagates downstream with an average velocity of $\sim 1.7 \text{ cm s}^{-1}$. The along-path length scale of these anomalies is $\sim 2,000 \text{ km}$ and their amplitude generally lies in the range $0.5\text{--}1.0^\circ\text{C}$, reaching a maximum between 5,000 and 6,000 km downstream of the tip of Florida at approximately 35° W, 50° N. As noted by Hansen and Bezdek⁴, this propagation speed is considerably slower than the near-surface currents in the core of the Gulf Stream ($>1 \text{ m s}^{-1}$ close to Cape Hatteras¹⁷). If, as seems likely, these signals are advective their speed must be determined by the weaker currents either to one side of the core or at some depth below the surface. Subsurface observations close to the path of the Gulf Stream/NAC show evidence of decadal

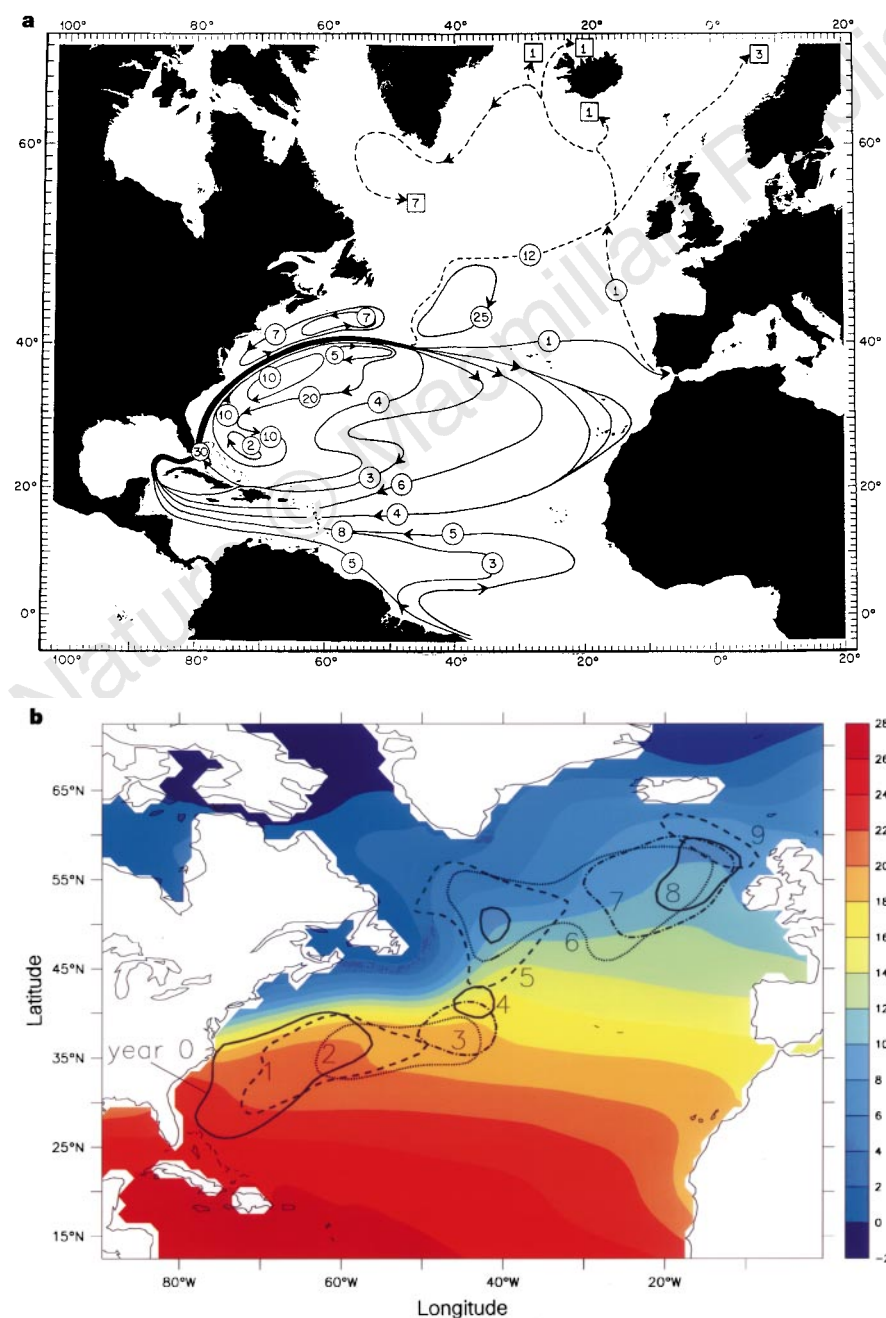


Figure 1 a, The upper-ocean circulation in the North Atlantic according to ref. 16. The lines indicate the mean flow of waters warmer than 7°C , and the circles numbers indicate the transport in sverdrups ($1 \text{ sverdrup} = 10^6 \text{ m}^3 \text{ s}^{-1}$). The main feature is the anticyclonic Subtropical Gyre, shown in solid lines. The cyclonic Subpolar Gyre lies to the north and is shown by dashed lines. The Gulf Stream lies at the western edge of the Subtropical Gyre. It flows polewards along the coast of North America before leaving the coast at Cape Hatteras. In mid-ocean the Gulf Stream splits into a part which recirculates southwards and another part—the North Atlantic Current—which flows northwards, and then north-eastwards towards Scotland. **b**, Correlation between low-frequency fluctuations in local wintertime SST and low-frequency fluctuations in wintertime SST averaged over the region $80\text{--}60^\circ \text{W}$, $31.5\text{--}38.5^\circ \text{N}$ (the vicinity of Cape Hatteras, VCH) as a function of lag. The contours pick out the regions where lag-correlation with VCH is maximized. The numbers next to each contour indicate the lag in years. In all cases VCH SST leads local SST. The contour value is 0.8 for lags of 0–8 years and 0.75 for the lag of 9 years. The contours are superimposed on the SST field averaged over all winters between 1945 and 1989 (colour scale in $^\circ\text{C}$). Monthly mean SST data were obtained from the NOAA SMD94¹⁵ analysis of ship observations for the period 1945–89 on a $1^\circ \times 1^\circ$ grid. The long-term mean annual cycle was removed from all the data. Winter-mean anomalies for each year are then computed by averaging the SST anomalies for the months November–April, and a five-year running mean is applied to the winter-mean anomalies. Correlations are only computed in the latitude band $25\text{--}65^\circ \text{N}$.

temperature fluctuations, possibly extending below the seasonal thermocline^{6,13,18}, and subsurface propagation in the Subpolar Gyre has also been reported^{19,20}. The existence of a deep signal, away from the regions of maximum shear and surface influence, may help to explain the remarkable coherence of the surface signature.

The SST power spectrum averaged along the Gulf Stream/NAC path exhibits a spectral peak at 12–14 years which exceeds the 98th percentile of the red-noise background that is expected if SST fluctuations arise as a local response to atmospheric variability²¹. In addition to this 12–14-year timescale, the prominence of the significant warm event which began in the mid-1940s and the significant cold event which began in the late 1960s suggests that lower-frequency, multidecadal, fluctuations² may also be associated with propagation along the Gulf Stream/NAC. Both timescales contribute to the correlations observed in Fig. 1b, and we stress the difficulty of distinguishing, with a short data record, a multi-decadal timescale from amplitude modulation of the decadal signals.

The 12–14-year timescale which we have identified with advection along the Gulf Stream/NAC agrees with that which previous authors^{1,7} have associated with a hypothetical coupled ocean–atmosphere mode. For any such coupled mode to exist, there must be a consistent atmospheric response to SST anomalies. How the atmosphere responds to mid-latitude SST is poorly under-

stood^{7–10}, but changes in the frequency, intensity or spatial distribution of mid-latitude storms are likely to be important^{8,10}. We may, therefore, expect particular sensitivity to SST anomalies in regions where these storms form. An especially important area is located along the southeastern coast of the USA between Cape Hatteras and Florida²² (which we call the ‘storm-formation region’) where severe wintertime land/sea temperature contrast fuels storm development through baroclinic instability. Figure 2 shows that SST in this region exhibits the decadal-timescale fluctuation, and furthermore that there is a strong correlation between these SSTs and a dipole-like fluctuation in North Atlantic sea-level pressure similar to that discussed in ref. 1. A warming of SST off the US coast, which we would expect to intensify the land/sea contrast and may enhance evaporation, is associated with a strengthening of the mean westerly winds in the mid-Atlantic especially in the region 50–30°W, 30–50°N. Although causality cannot reliably be inferred from correlation alone, this relationship is in line with the theoretical response of the mean flow to an intensified storm track²³ and therefore supports a hypothesis that the sea-level pressure dipole is a response to SST in the storm-formation region.

Increased storminess and/or mean winds will cool SSTs through local mixed layer and Ekman layer processes. Wind and net surface heat flux fields (not shown) indicate that cooling of SSTs in the vicinity of the NAC (40–55°N) and in the tropical North Atlantic

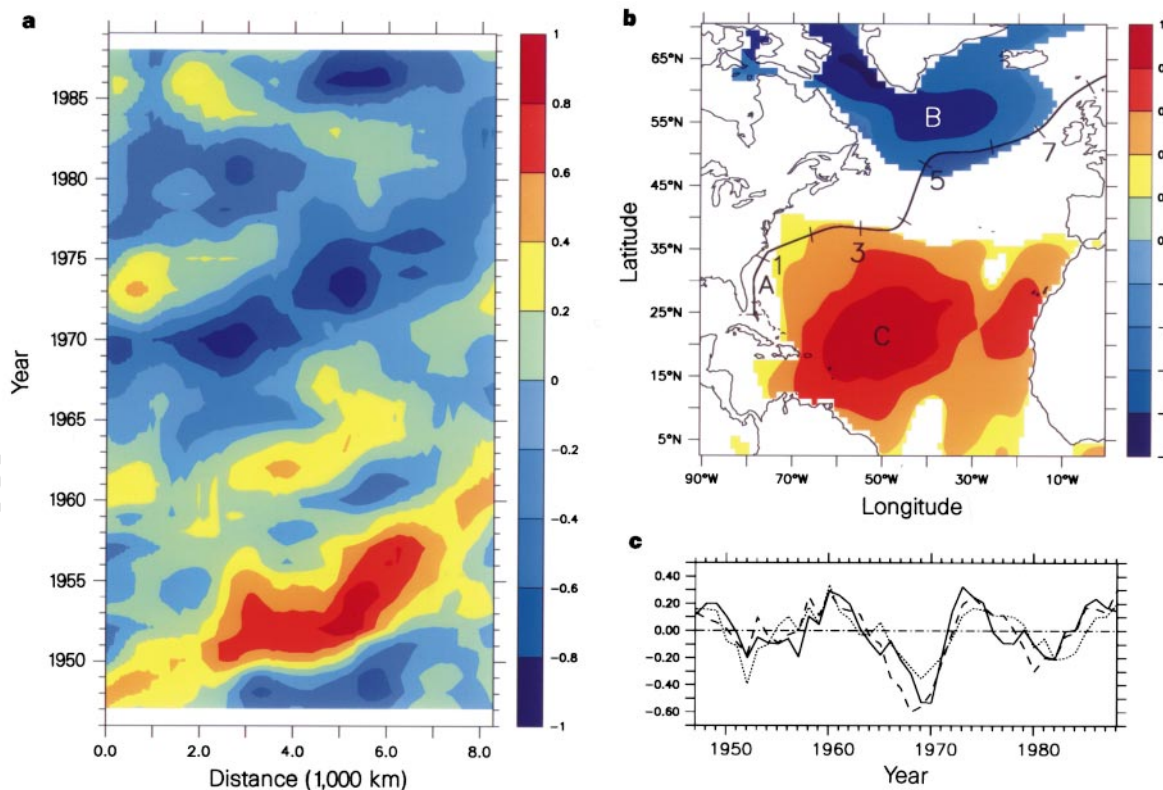


Figure 2 a, SST anomalies (colour scale in °C) as a function of time and distance along the path of the Gulf Stream/North Atlantic Current. Winter (November–April) anomalies only are used, smoothed with a three-year running mean. The path along which the plot is computed (shown in **b** as a solid black line) was taken from ref. 16 (Fig. 1a). The horizontal axis gives distance in thousands of kilometres measured along the path from the start point at 78.5°W, 24.5°N near the tip of Florida. **b**, The fraction of the low-frequency variability (root-mean-squared anomalies) in winter-mean sea-level pressure that can be accounted for by a linear response to SST in the storm-formation region (SFR; 82–69°W, 25–35°N, marked with a letter A). Positive (negative) values indicate fluctuations (anti-) correlated with SFR SST. In **c**, the solid line is a time series of winter-mean SST in region SFR, and the vertical scale is in °C. The dashed line shows winter-mean

sea-level pressure in the region 50–30°W, 50–60°N (letter B in panel **b**), multiplied by –1. The dotted line shows winter-mean sea-level pressure in the region 60–40°W, 15–25°N (letter C in panel **b**). Both the sea-level pressure time series are renormalized to have approximately the same variance as the time series of SST, and all three time series are detrended and smoothed with a three-year running mean. Before performing the regression for panel **b** both the SST and sea-level pressure time series are binned into 2-year winter means. The SFR SST time series only is further smoothed with a 3-point running mean. Sea-level pressure at any given location is then regressed on SFR SST, a constant term, and a trend, with the residual modelled as an AR(1) (‘red noise’) process²⁶. Values not significant at the 1.5 standard deviation level are not plotted.

(5–20° N) is associated with warm SST in the storm-formation region. Thus fluctuations in SST off the southeastern coast of the USA may affect remote regions of the Atlantic both quasi-instantaneously, via this atmospheric teleconnection, and after a delay of some years via advective propagation through the ocean.

We examine how oceanic advection and atmospheric teleconnections may combine to determine the decadal cycle in SST by regressing SST and sea-level pressure at various lags onto SST in the storm-formation region (Fig. 3). The analysis is designed (see Fig. 3 legend) to focus on the decadal timescale. The regression map at lag zero shows that warm SST in the storm formation region is associated with cold SST in the NAC region and in the tropical North Atlantic. These cold SSTs arise largely as the local SST response to heat-flux anomalies¹. We suggest that these heat flux anomalies arise as part of the atmospheric response to SST in the storm-formation region. At lags of 2 and 4 years, consistent with the reduced amplitude of SST anomalies in this region, the sea-level pressure dipole is much weaker than at lag 0. The SST regression map evolves in a manner broadly consistent with our expectations of advection (Fig. 1a). The warm anomaly in the storm-formation region moves eastwards and then appears to split into two branches over the Mid-Atlantic Ridge, in good agreement with the bifurcation of the Gulf Stream indicated in Fig. 1a. The cold anomaly in the North Atlantic moves eastwards along the path of the NAC. At a lag of 4 years we see the development of cold anomalies in the Gulf of Mexico which, at a lag of 6 years, spread into the storm-formation region. The development of cold anomalies in the latter region stimulates, according to our hypothesis, an atmospheric response of

the opposite sign to that with which we began. This atmospheric response forces warm anomalies in the NAC region and in the tropical North Atlantic, developing a pattern which is 180° out of phase with that seen at lag zero.

The explanation for a preferred timescale is likely to involve feedbacks to SST in the storm-formation region, mediated by the quasi-instantaneous atmospheric response and by subsequent oceanic processes that introduce a lag. The effects of the atmospheric fluctuations on the oceanic mixed layer in the tropical Atlantic, and in the vicinity of the Gulf Stream front (where ventilation may be affected), could be fed back to the storm-formation region by advection. Alternatively, fluctuations in the wind-stress curl over the Subtropical Gyre could generate fluctuations in the strength of the gyre circulation²⁴. These fluctuations could also influence SST in the storm-formation region, but in this case the dominant timescale is associated with the propagation of baroclinic Rossby waves. The fact that anomalies arise in the Gulf of Mexico before their arrival in the storm-formation region is consistent with both mechanisms. In reality it is likely that both advection and wave propagation influence the timescale²; external forcing might also play a role²⁵.

In the NAC region and in the tropical North Atlantic, interference between the instantaneous atmospheric influence and the delayed advective influence of SST in the storm-formation region could also favour the emergence of a preferred timescale³¹, tending to suppress (amplify) variability on timescales that are odd (even) multiples of the ~6-year (Fig. 1b) advection time from the storm-formation region to these regions. Both the instantaneous and delayed signals will have contributed to the correlations seen in Fig. 1b. We note

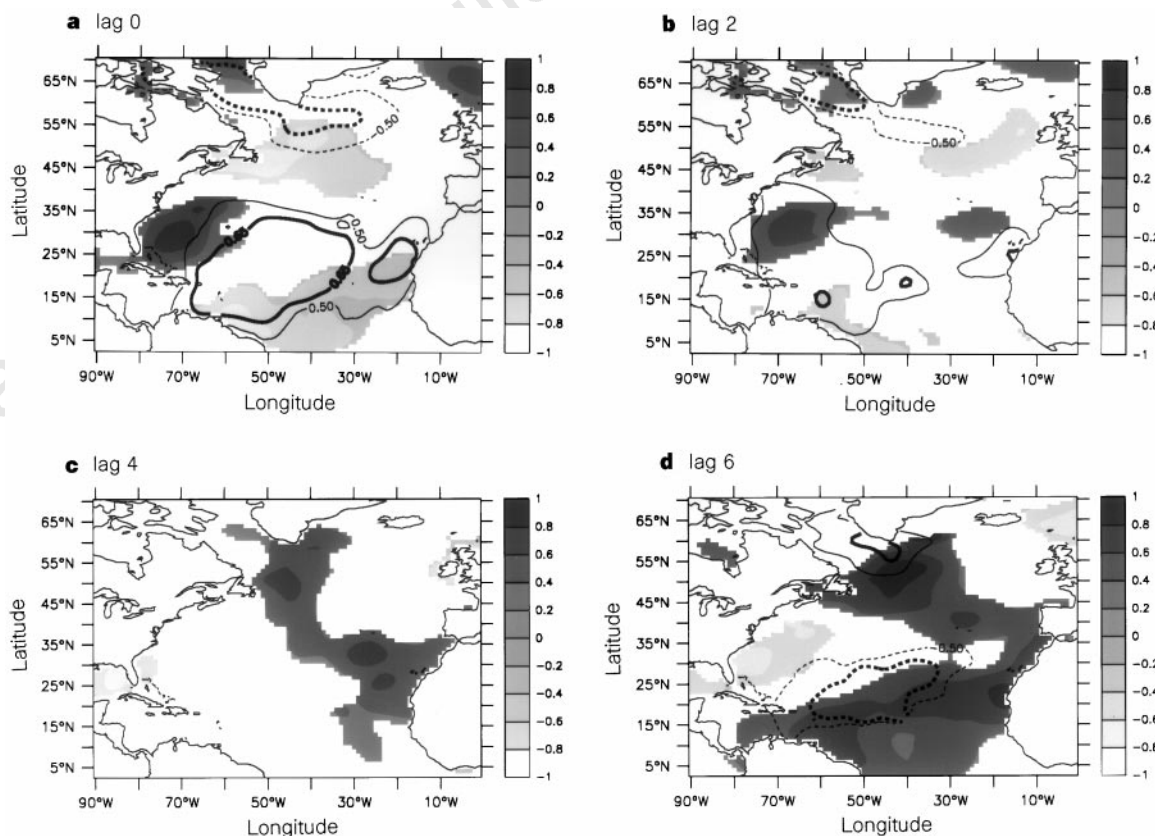


Figure 3 The fraction of low-frequency local SST variability (grey-scale) and low-frequency local sea-level pressure variability (contours) accounted for by a lagged version of the low-frequency SST variability in the storm-formation region, SFR; see Fig. 2 legend. SST in region SFR leads local SST, and local sea-level pressure, by the number of years indicated on each panel. The regression analysis is identical to that described in Fig. 2 legend but with winter-mean SST

replacing winter-mean sea-level pressure for the grey-scale maps, and with varying lag. The contour levels shown for the sea-level pressure regression are ± 0.5 (thin lines) and ± 0.65 (thick lines). Dashed lines indicate negative correlations. In **c**, the fraction of sea-level pressure variability accounted for fails to reach the 0.5 level, so no contours appear.

that our account of the decadal variability in tropical North Atlantic SST²⁶ is an alternative to the hypothesis that this region is under the control of processes local to the tropical Atlantic^{27,28}.

By exploiting the memory inherent in the decadal fluctuations, our results suggest it may be possible to predict a significant fraction of the low-frequency variability in North Atlantic SST and sea-level pressure several years in advance. Although further work will be needed before a serious attempt at forecasting can be made, particularly investigation with models of the mechanisms we have discussed, we consider briefly the outlook for the decadal fluctuations. We have examined the most recent evolution of North Atlantic SST in the NOAA NCEP analyses²⁹ and of North Atlantic sea-level pressure in the NCEP reanalyses³⁰. SST in the storm-formation region reached a maximum in the late 1980s and has since been falling. In line with our expectations of the atmospheric response, sea-level pressure in the region marked C in Fig. 2b peaked at the end of the 1980s and has since been falling, whereas sea-level pressure in the region marked B reached a minimum at the same time and has since been rising. Also consistent with our understanding, SST in the mid-Atlantic at ~50° N reached a minimum in the late 1980s and has since been rising. Similar, if somewhat less clear, behaviour is seen in tropical North Atlantic SST. Based on the apparent period of 12–14 years, and to the extent that other influences (for example, El Niño and multidecadal fluctuations) allow, we expect all these trends to continue into the next century before reversing. □

Received 12 May; accepted 17 June 1997.

- Deser, C. & Blackmon, M. L. Surface climate variations over the North Atlantic ocean during winter: 1900–1989. *J. Clim.* **6**, 1743–1753 (1993).
- Kushnir, Y. Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *J. Clim.* **7**, 141–157 (1994).
- Battisti, D. S., Bhatt, U. S. & Alexander, M. A. A modelling study of the interannual variability in the wintertime North Atlantic ocean. *J. Clim.* **8**, 3067–3083 (1995).
- Hansen, D. V. & Bezdek, H. F. On the nature of decadal anomalies in North Atlantic sea surface temperature. *J. Geophys. Res.* **101**, 8749–8758 (1996).
- Delworth, T. L. North-atlantic interannual variability in a coupled ocean-atmosphere model. *J. Clim.* **9**, 2356–2375 (1996).
- Molinari, R. L., Mayer, D., Festa, F. & Bezdek, H. Multiyear variability in the near surface temperature structure of the midlatitude western North Atlantic ocean. *J. Geophys. Res.* **102**, 3267–3278 (1997).
- Grötzner, A., Latif, M. & Barnett, T. P. A decadal climate cycle in the North Atlantic ocean as simulated by the ECHO Coupled GCM. *J. Clim.* (submitted).
- Palmer, T. N. & Sun, Z. A modeling and observational study of the relationship between sea surface temperature in the Northwest Atlantic and the atmospheric general circulation. *Q. J. R. Meteorol. Soc.* **111**, 947–975 (1985).
- Peng, S., Mysak, L. A., Ritchie, H., Derome, J. & Dugas, B. The differences between early and midwinter atmospheric responses to sea surface temperature anomalies in the Northwest Atlantic. *J. Clim.* **8**, 137–157 (1995).
- Kushnir, Y. & Held, I. M. Equilibrium atmospheric response to North Atlantic SST anomalies. *J. Clim.* **9**, 1208–1220 (1996).
- Griffies, S. M. & Bryan, K. Predictability of north atlantic multidecadal climate variability. *Science* **275**, 181–184 (1997).
- Mann, M. E. & Park, J. Joint spatiotemporal modes of surface temperature and sea level pressure variability in the Northern Hemisphere during the last century. *J. Clim.* **9**, 2137–2161 (1996).
- Houghton, R. W. Subsurface quasi-decadal fluctuations in the North Atlantic. *J. Clim.* **9**, 1361–1373 (1996).
- Bjerknes, J. Atlantic air-sea interaction. *Adv. Geophys.* **10**, 1–82 (1964).
- daSilva, A., Young, A. C. & Levitus, S. NOAA SMD94 (Tech. Rep. 6, US Dept of Commerce, NOAA, NESDIS, Washington DC, 1994).
- Schmitz, W. J. & McCartney, M. S. On the North Atlantic circulation. *Rev. Geophys.* **31**, 29–49 (1993).
- Stommel, H. *The Gulf Stream* (Univ. California Press, Berkeley, 1965).
- Levitus, S., Antonov, J. I. & Boyer, T. P. Interannual variability of temperature at a depth of 125 metres in the North Atlantic ocean. *Science* **266**, 96–99 (1994).
- Reverdin, G., Cayan, D. & Kushnir, Y. Decadal variability of hydrography in the upper North Atlantic 1948–1990. *J. Geophys. Res.* **102**, 8505–8531 (1997).
- Dickson, R. R., Meincke, J., Malmberg, S.-A. & Lee, A. J. The “Great Salinity Anomaly” in the northern North Atlantic 1968–1982. *Prog. Oceanogr.* **20**, 103–151 (1988).
- Frankignoul, C. Sea surface temperature anomalies, planetary waves, and air-sea feedback in the middle latitudes. *Rev. Geophys.* **23**, 357–390 (1985).
- Roebber, P. J. Statistical analysis and updated climatology of explosive cyclones. *Mon. Weath. Rev.* **112**, 1577–1589 (1984).
- Hoskins, B. J. in *Large Scale Dynamical Processes in the Atmosphere* (eds Hoskins, B. J. & Pearce, R. P.) 169–199 (Academic, London, 1983).
- Latif, M. & Barnett, T. P. Causes of decadal climate variability over the North Pacific/North American sector. *Science* **266**, 634–637 (1994).
- White, W. B., Lean, J., Cayan, D. R. & Dettinger, M. D. Response of global upper ocean temperature to changing solar irradiance. *J. Geophys. Res.* **102**, 3255–3266 (1997).
- Allen, M. R. & Smith, L. A. Investigating the origins and significance of low-frequency modes of climate variability. *Geophys. Res. Lett.* **21**, 883–886 (1994).
- Mehta, V. M. & Delworth, T. Decadal variability of the tropical Atlantic ocean surface temperature in shipboard measurements and in a global ocean-atmosphere model. *J. Clim.* **8**, 172–190 (1995).
- Chang, P., Ji, L. & Li, H. A decadal climate variation in the tropical Atlantic ocean from thermodynamic air-sea interactions. *Nature* **385**, 516–518 (1997).

- Reynolds, R. W. & Smith, T. M. Improved sea surface temperature analyses using optimum interpolation. *J. Clim.* **7**, 929–948 (1994).
- Kalnay, E. et al. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**, 437–471 (1996).
- Saravanan, R. & McWilliams, J. C. Advective ocean-atmosphere interaction: an analytical stochastic model with implications for decadal variability. *J. Clim.* (in the press).

Acknowledgements. We thank T. Haine for comments on earlier drafts, and W. Schmitz and M. McCartney for permission to reproduce Fig. 1a. R.T.S. was supported by the UK NERC Universities Global Atmospheric Modelling Programme, M.R.A. was supported by the NERC/RAL Ocean Dynamics Service Level Agreement.

Correspondence should be addressed to R.T.S. (e-mail: atmrtm@atm.ox.ac.uk).

Orbitally paced climate oscillations across the Oligocene/Miocene boundary

James C. Zachos*, Benjamin P. Flower†‡ & Hilary Paul*

* Earth Science Department, † Institute of Marine Sciences, University of California, Santa Cruz, California 95064, USA

The late Oligocene and early Miocene periods, some 21 to 27 million years ago, have generally been viewed as times of moderate global warmth and ice-free conditions. Yet several lines of evidence suggest that this interval was punctuated by at least one, and possibly several, episodes of high-latitude cooling and continental glaciation^{1–3}. Here, we present stable-isotope and per cent coarse-fraction data from an equatorial, western Atlantic deep-sea-sediment core that provide high-resolution records of the climate variability across the Oligocene/Miocene transition (22.5–25.7 million years ago). A strong 40-kyr periodicity in the oxygen isotope record is consistent with a high-latitude orbital (obliquity) control on ice-volume and temperature. Orbital influences are also apparent at precession and eccentricity frequencies, including a series of ~400-kyr oscillations that culminate in distinct maxima at the Oligocene/Miocene boundary, about 23.7 million years ago. Covariance between the carbon and oxygen isotope records suggests that the oceanic carbon cycle may have contributed to global cooling during the ~400-kyr cycles, particularly at the Oligocene/Miocene boundary.

The Oligocene/Miocene boundary represents a significant climate transition in Cenozoic Earth history; ocean temperatures were slowly increasing and continental ice-volume was apparently decreasing from the large accumulations of the middle–late Oligocene^{2–4}. Oxygen isotope records indicate that this trend may have been interrupted by a series of brief glaciations that began in the earliest Miocene and continued up to the time of significant middle Miocene cooling². The first and largest of these events, Mi-1, occurred at the Oligocene/Miocene boundary.

As with most other pre-Pliocene intervals, previous isotope records that span the Oligocene/Miocene boundary lack the sampling density needed to resolve palaeoenvironmental changes occurring on timescales of <10⁵ years. Attempts to construct higher resolution records have been hampered either by the lack of continuous, undisturbed sedimentary sections, or by diagenetic overprinting. These deficiencies were recently resolved with the successful recovery of an expanded and hiatus-free section of upper Oligocene and lower Miocene nannofossil chalk in Hole 929A, on Ceara rise in the western equatorial Atlantic Ocean (ref. 5) (4360 m water depth). This section is characterized by prominent decimetre-scale oscillations in sediment colour and magnetic susceptibility with periods similar to those associated with orbital forcing⁶.

‡ Present address: Department of Marine Science, University of South Florida, St Petersburg, Florida 33701, USA.

Using samples from Hole 929A, we assembled high-resolution (10-cm sampling) isotope and per cent coarse-fraction time series for the Oligocene/Miocene boundary (22.5–25.7 Myr ago). Initially, samples were collected from four cores (929A 34X–37X) spanning a 40-m interval across the boundary⁷. We subsequently increased the length of the record to 60 m by adding samples from two additional

cores (929A 33X–38X) to create time series nearly 3.2 Myr long across the Oligocene/Miocene boundary. Core recovery was 100% over this interval. Site-to-site comparisons of magnetic susceptibility and colour reflectance data suggest inter-core gaps of <40 cm (~20 kyr)⁵. The stable-isotope record, plotted in Fig. 1 along with per cent coarse fraction (% CF), is based on the carbon and oxygen isotope ratios of the benthic foraminifer, *Cibicides mundulus*. The age model is based on sedimentation rates derived from the Hole 929A biostratigraphy⁵ and the wavelength of the 40-kyr cycle in the magnetic susceptibility record⁶ (Table 1). Ages are cumulative, starting from an assigned age of 22.5 Myr for the uppermost

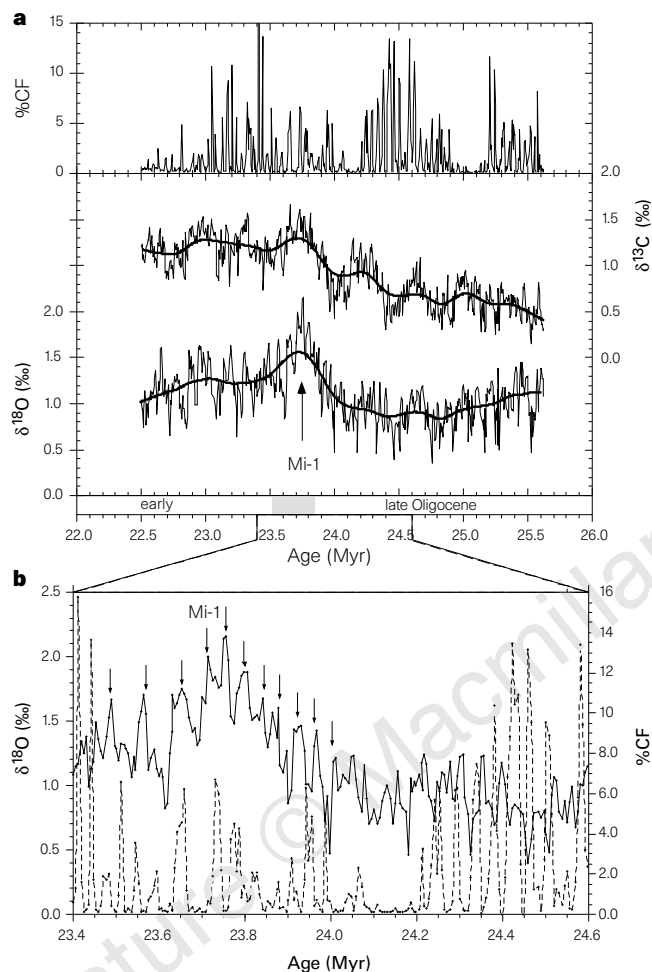


Figure 1 **a**, Benthic foraminifer stable-isotope and % coarse fraction (%CF) records for Hole 929A plotted versus age. Each sample was freeze-dried, weighed, washed and wet sieved over a 63- μm screen. The >63 μm fraction was dried and weighed to calculate %CF. From five to ten specimens of benthic foraminifer taxa, *Cibicides mundulus*, were then picked from the >150- μm fraction of each sample and analysed for carbon and oxygen isotopic composition. Stable-isotope analyses were carried out in the Stable Isotope Laboratory at the University of California, Santa Cruz using a common acid-bath device coupled to either a Fisons PRISM or an OPTIMA mass spectrometer¹². All values are reported in the per mil (‰) notation relative to Pee Dee Belemnite standard. External precision based on duplicates and triplicates is better than ± 0.09 for both C and O isotopes. The %CF represents the mass of the >63- μm size fraction relative to the total mass of each sample. The heavy lines in the stable-isotope plots represent the curve fits (locally weighted (15%) means) used for detrending. The age model is based on sedimentation rates inferred from the wavelength of the 40-kyr cycle in magnetic susceptibility⁶. This age model proved to be superior to one based exclusively on biostratigraphic datums which, because of dissolution and low abundances of planktonic microfossils at this site, tended to be less reliable. Nevertheless, the differences between the two age models are relatively small (see Supplementary Information). **b**, Expanded view of the oxygen-isotope (solid line; left-hand ordinate) and %CF (dashed line; right-hand ordinate) records for the period 23.4–24.6 Myr ago. The arrows denote individual "glacial maxima" associated with the Mi-1 event.

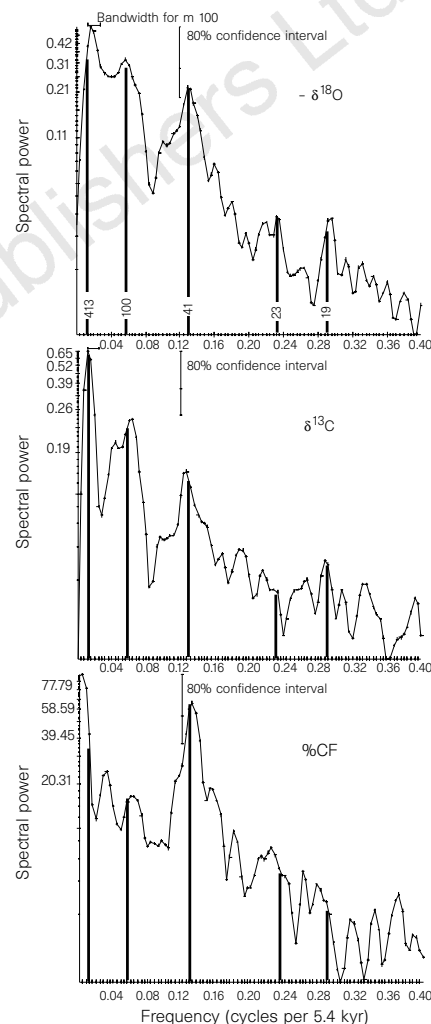


Figure 2 Power spectra (variance versus frequency) of three variables, the benthic $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$, and %CF values. The left-hand ordinate shows the spectral power of each record. Bandwidth (0.0133) and confidence intervals (80%) are shown. The time-series analysis was carried out using a numerical approach based on the Blackman–Tukey method (SPECTX; Oregon State Univ.). Following SPECMAP convention, the oxygen isotope record was multiplied by -1 so that interglacials (less ice and warmer temperatures) are represented by higher values. After converting to age, the data set was resampled at an equal time interval of 5.4 kyr ($n = 578$). The number of lags (m) was set to 100. To resolve the lower-frequency components of the oxygen and carbon isotope time series, locally weighted means (Fig. 1a) were fitted to the records and subtracted to remove the longer-term trends from each time series. Residual long-term trends were then removed using a linear detrend. This pre-processing substantially improved the degree to which variance at the low-frequency (~400 and 100 kyr) end of the spectrum could be resolved. The vertical bars show the frequencies of the primary orbital periods.

level (core 929A-33X, 10 cm; 301.10 m below sea floor) of the sampled section. This age is based on an interpolation of two biostratigraphic age datums calibrated against the most recent Geomagnetic Polarity Time Scale (GPTS)⁸.

The oxygen-isotope time series shows changes on several different scales, including pervasive high- and low-frequency ($\sim 10^4$ – 10^5 yr) oscillations and a brief positive $\delta^{18}\text{O}$ excursion of 1.2‰ just above the Oligocene/Miocene boundary (Fig. 1a). The high-frequency oscillations have amplitudes of between 0.5 and 0.6‰; the lower-frequency oscillations have slightly larger amplitudes. The 1.2‰ excursion initiated 24.0 Myr ago and peaked at 23.7 Myr. The magnitude and timing suggest it is equivalent to the Mi-1 event. In detail, it consisted of a series of higher-frequency cycles of varying amplitude superimposed on a longer-term increase (Fig. 1b). The carbon isotope record also shows variability on several scales, including high- and low-frequency oscillations, and a long-term rise. The high-frequency cycles are of small amplitude, generally <0.4 ‰. The lower-frequency cycles are of significantly larger amplitude, in excess of 0.8‰, and appear to covary with oxygen isotopes over most of the record. The long term rise began ~ 25.0 Myr ago and peaked at 23.65 Myr with mean $\delta^{13}\text{C}$ values increasing by >0.7 ‰.

The coarse fraction ($>63\text{-}\mu\text{m}$) of these sediments consists largely of planktonic foraminifer tests. The %CF record is dominated by high-frequency variability ($\sim 10^4$ yr) with values varying between 0.1 and 12%. A low-frequency pattern is present as well; there are three intervals, each 100–200 kyr long, with low mean concentrations and values of %CF rarely exceeding 1%. These three intervals are centred at roughly 22.8, 24.1 and 25.1 Myr ago.

To identify the main periods of variance in the isotope and %CF records, as well as internal phase relationships (coherency) between variables, spectral and cross-spectral analyses were carried out using a program based on the Blackman–Tukey method⁹ (for details see legends of Figs 2 and 3). High concentrations of variance in the $\delta^{18}\text{O}$

spectrum were identified at frequency components that correspond to the Milankovitch periods of 19, 23, 40, 100 and ~ 413 kyr (0.284, 0.235, 0.130, 0.054 and 0.013 cycles per 5.4 kyr; Fig. 2). The $\delta^{13}\text{C}$ spectra shows concentration of variance at the 19-, 41-, 100- and ~ 413 -kyr periods, whereas the %CF spectra is dominated by the 40-kyr period.

At the ~ 400 - and 100-kyr periods, $-\delta^{18}\text{O}$ ($\delta^{18}\text{O}$ is multiplied by -1 following SPECMAP convention) and $\delta^{13}\text{C}$ show a high-degree coherency ($k = 0.91$ and 0.89 , respectively), with $-\delta^{18}\text{O}$ (glacial minima) nearly 180° out of phase with $\delta^{13}\text{C}$ (Fig. 3a). These variables are also coherent ($k = 0.72$) and nearly in phase at the 40-kyr period. A high degree of coherency ($k = 0.85$) also occurs between %CF and $\delta^{18}\text{O}$ at the 40-kyr period with $-\delta^{18}\text{O}$ lagging %CF by 40° (4.4 ± 1.7 kyr; Fig. 3b).

The strong response in the oxygen isotope record at the obliquity (40-kyr) period is significant. Because the effects of obliquity on insolation are strongest at high latitudes¹⁰, oscillations with a 40-kyr period suggest a high-latitude climate control (ice-volume and/or temperature) on $\delta^{18}\text{O}$ across the Oligocene/Miocene boundary. This finding, along with evidence of similar periodicities in earliest Oligocene and early middle Miocene oxygen isotope records^{11,12}, as well as in Plio-Pleistocene records, suggests that a strong response to obliquity forcing has been a common feature of the global climate signal for at least the past 34 Myr.

These obliquity-paced climate oscillations were closely linked to deep-water circulation patterns and chemistry. At the 40-kyr period, %CF appears to be tightly coupled to $\delta^{18}\text{O}$ over most of this record (Fig. 3a). Three factors can influence %CF: surface production, dilution by terrigenous debris, and dissolution¹³. At Hole 929A, a general association of low %CF with increased fragmentation and low carbonate concentration (as inferred from magnetic susceptibility and colour reflectance⁵) indicates that dissolution was the controlling factor. This coupling of $\delta^{18}\text{O}$ and carbonate dissolution is similar to patterns observed in Quaternary

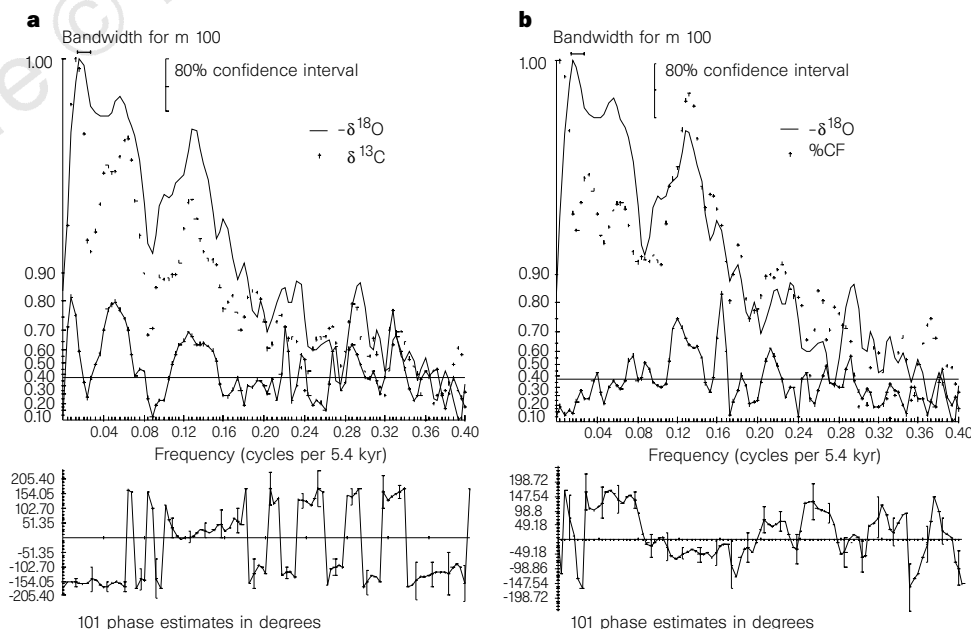


Figure 3 Top panels, coherence spectra of the $\delta^{13}\text{C}$ (panel a) and %CF (panel b) time series relative to $\delta^{18}\text{O}$ (solid lines). Following SPECMAP convention, all $\delta^{18}\text{O}$ values were multiplied by -1 so that phase relationships are relative to glacial minima rather than maxima. The isotope time series were detrended by subtracting a locally weighted mean, and then resampled at a constant 5.4-kyr time interval. Variance density and coherency (solid line with plus signs) were

calculated based on 100 lags of the autocovariance function. The coherencies are plotted on a hyperbolic arctangent scale along the y-axis so that the confidence intervals for the coherency estimates are a constant length. The horizontal line is the 80% confidence interval testing the non-zero hypothesis for the estimates. Bottom panels, the phase angles (in degrees) and associated 80% confidence intervals (vertical lines).

Table 1 Age model parameters

Cores	Depth (m below sea floor)	Wavelength* (m)	Sed. rate* (m Myr ⁻¹)	Age (Myr)
929A-33-34X	301.05–320.36	0.683	171	22.500–23.626
929A-35-36X	320.45–339.61	0.659	16.5	23.630–24.793
929A-37-38X	339.70–356.80	0.826	20.7	24.797–25.623

* From ref. 6: 'Wavelength' means wavelength of the primary cycles of magnetic susceptibility, 'sed. rate' means sedimentation rate.

sediments^{14,15} and can be explained by increased flux of more corrosive bottom water into the equatorial Atlantic during glacial periods. The presence of a south-to-north decrease in deep-water $\delta^{13}\text{C}$ during the early Miocene¹⁶ suggests a Southern Ocean source for this more corrosive water.

The Hole 929 $\delta^{18}\text{O}$ record resolves two critical features of the Mi-1 glacial maximum. The first is the structure of the Mi-1 excursion; it actually consisted of a series of higher-frequency (40-, 100-kyr) oscillations superimposed on a long-term increase (Fig. 1b). The direct response at the 100-kyr cycle appears to be slightly stronger just after the Mi-1 event, although further work is needed to resolve the detailed evolution of the climate spectra over this and other intervals. The second important feature is the magnitude of Mi-1; a 1.2‰ increase requires an accumulation of continental ice equivalent in volume to the present-day East Antarctic Ice Sheet, or 5–6°C cooling of bottom waters, or, more likely, some combination of ice-volume/temperature change. The presence of Upper Oligocene and Lower Miocene glacial marine sediments in the Ross Sea¹⁷, as well as new evidence that links a prominent eustatic lowering¹⁸ to Mi-1¹⁹ suggests that part of the $\delta^{18}\text{O}$ increase was due to ice-sheet expansion.

Perhaps the most intriguing features of the isotope spectra are the high-amplitude ~400-kyr oscillations, particularly in $\delta^{13}\text{C}$. Variance in climate near this period, as well as at the 100-kyr period, is thought to originate from an asymmetrical response mechanism that preferentially introduces variance from the warmer portions of the eccentricity modulated precession cycle²⁰. Although rarely recorded in Pliocene/Pleistocene^{20,21}, ~400-kyr oscillations commonly arise in lower-resolution Oligocene and Miocene benthic foraminiferal $\delta^{13}\text{C}$ time series^{12,22–24}. Such a strong expression of this period in the carbon isotope record indicates a nonlinear relation between the carbon cycle and eccentricity forcing. At Hole 929A, $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ are closely coupled at the ~400-kyr period (lag $\sim 24 \pm 20$ kyr) as well as at 100-kyr period, with $\delta^{18}\text{O}$ leading $\delta^{13}\text{C}$. Covariance between oxygen and carbon in early Oligocene and middle Miocene records has previously been attributed to the effects of climate change on ocean/atmosphere circulation, ocean productivity, and organic carbon^{23–25}. Investigators speculate that as high-latitude temperatures decrease and ice sheets grow, oceanic and atmospheric circulation intensify, increasing oceanic turnover, productivity and organic-carbon fluxes. In turn, burial rates of organic carbon increase thereby driving mean ocean $\delta^{13}\text{C}$ towards higher values²⁶. The process apparently reverses as climate warms and/or nutrient levels are reduced. This coupling may have helped amplify the response at the lower frequencies.

Close coupling of climate (as expressed by $\delta^{18}\text{O}$) and the oceanic carbon cycle may also explain the unusual scale of the Mi-1 event. If the low-frequency carbon-isotope cycles reflect periodic changes in the rate of organic-carbon burial as postulated above, it is likely that the atmospheric partial pressure of CO_2 (p_{CO_2}) was oscillating as well, and generating a positive feedback. We speculate that the variations in p_{CO_2} during the Oligocene were within a range where the feedback on the global radiative balance was comparatively small until the earliest Miocene (Mi-1) when p_{CO_2} had dropped below some threshold. The process of reaching this threshold was probably gradual, as suggested by the progressive step by step increase in carbon isotope ratios before Mi-1, a trend which

indicates significant oceanic carbon-cycle change during the latest Oligocene.

The Hole 929A isotope times series presented here provide critical new insight into the character of climate variability across the Oligocene/Miocene transition. The persistence of a strong 40-kyr period in the $\delta^{18}\text{O}$ record is consistent with the presence of Antarctic ice sheets during this period of moderate global warmth. Documentation of 100- and 400-kyr periods provides strong support for a palaeoceanographic response to the longer-period Milankovitch cycles. Covariance between $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ at these eccentricity periods hints at a strong coupling between climate and the ocean carbon cycle. Tuning of the Hole 929 isotope time series to orbital spectra (N. J. Shackleton, personal communication) will ultimately produce a standard reference section that will provide a degree of chronostratigraphic control for the late Oligocene to early Miocene that approaches that of the Quaternary. □

Received 10 March; accepted 18 June 1997.

1. Leckie, R. M. & Webb, P. N. Late Oligocene-early Miocene glacial record of the Ross Sea, Antarctica: evidence from DSDP Site 270. *Geology* **11**, 578–582 (1983).
2. Miller, K. G., Wright, J. D. & Fairbanks, R. G. Unlocking the ice house: Oligocene-Miocene oxygen isotopes, eustasy and margin erosion. *J. Geophys. Res.* **96**, 6829–6848 (1991).
3. Wright, J. D. & Miller, K. G. Miocene stable isotope stratigraphy, Site 747, Kerguelen Plateau. *Proc. ODP Sci. Res.* **120**, 855–866 (1992).
4. Stott, L. D., Kennett, J. P., Shackleton, N. J. & Corfield, R. M. The evolution of Antarctic surface waters during the Paleogene, inferences from the stable isotopic composition of planktonic foraminifera. *Proc. ODP Sci. Res.* **113**, 849–864 (1990).
5. Curry, W. B. et al. *Init. Rep. ODP A* **154**, (1995).
6. Weedon, G. P., Shackleton, N. J. & Pearson, P. N. The Oligocene time scale and cyclostratigraphy on the Ceara Rise, Western Equatorial Atlantic. *Proc. ODP Sci. Res.* (in the press).
7. Flower, B. P., Zachos, J. C. & Paul, H. Milankovitch-scale variability recorded near the Oligocene/Miocene Boundary: Hole 929A. *Proc. ODP Sci. Res.* (in the press).
8. Berggren, W. A. et al. in *Geochronology, Time Scales and Global Stratigraphic Correlation* (eds Berggren, W. A., Kent, D. V., Aubry, M.-P. & Hardenbol, J.) 129–212 (Spec. Publ. 54, Soc. Econ. Paleontol. Mineral., Tulsa, OK, 1995).
9. Jenkins, G. M. & Watts, D. G. *Spectral Analysis and its Applications* (Holden-Day, San Francisco, 1968).
10. Berger, A. & Loutre, M. F. Insolation values for the climate of the last 10 m.y. *Quat. Sci. Rev.* **10**, 297–317 (1991).
11. Pisias, N. G., Shackleton, N. J. & Hall, M. A. Stable isotope and calcium carbonate records from hydraulic piston cored Hole 574A: high-resolution records from the middle Miocene. *Init. Repts DSDP* **85**, 735–748 (1985).
12. Zachos, J. C., Quinn, T. M. & Salamy, K. Earliest Oligocene climate transition: Constraints from high resolution (10^4 yr) deep-sea foraminiferal $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ time series. *Paleoceanography* **11**, 251–266 (1996).
13. Wu, G., Herguera, J. C. & Berger, W. H. Differential dissolution, modification of late Pleistocene oxygen isotope records in the western equatorial Pacific. *Paleoceanography* **5**, 581–594 (1990).
14. Bassinot, F. C. et al. Coarse fraction fluctuations in pelagic carbonate sediments from the tropical Indian Ocean: A 1500-kyr record of carbonate dissolution. *Paleoceanography* **9**, 579–600 (1994).
15. Curry, W. B., Duplessy, J. C., Labeyrie, L. D., Oppo, D. & Kallel, N. P. Changes in the distribution of $\delta^{13}\text{C}$ of ECO_2 between the last glaciation and the Holocene. *Paleoceanography* **3**, 317–342 (1988).
16. Wright, J. D. & Miller, K. G. in *The Antarctic Paleoenvironment: A Perspective on Global Change* (eds Kennett, J. P. & Warnke, D. A.) 141–165 (Antarctic Res. Ser. 56, Am. Geophys. Un., Washington DC, 1993).
17. Barrett, P. J. (ed.) *Antarctic Cenozoic History From CIROS-1 Drill Hole McMurdo Sound*. *NZ DSIR Bull.* **245**, 1–251 (1989).
18. Haq, B. U., Hardenbol, J. & Vail, P. R. Chronology of fluctuating sea levels since the Triassic. *Science* **235**, 1156–1167 (1987).
19. Miller, K. G. et al. Drilling and dating New Jersey Oligocene-Miocene sequences; ice volume, global sea level, and Exxon records. *Science* **271**, 1092–1095 (1996).
20. Clemens, S. C. & Tiedemann, R. Eccentricity forcing of Pliocene–Early Pleistocene climate revealed in a marine oxygen isotope record. *Nature* **385**, 801–804 (1997).
21. Mix, A. C. et al. Benthic Foraminifer stable isotope record from Site 849, 0–5 Ma: Local and global climate changes. *Proc. ODP Sci. Res.* **138**, 371–412 (1995).
22. Diester-Haass, L. & Zahn, R. Eocene-Oligocene transition in the Southern Ocean: History of water mass circulation and biological productivity. *Geology* **24**, 163–166 (1996).
23. Woodruff, F. & Savin, S. M. Mid-Miocene isotope stratigraphy in the deep sea: high-resolution correlations, paleoclimatic cycles, and sediment preservation. *Paleoceanography* **6**, 755–806 (1991).
24. Flower, B. P. & Kennett, J. P. Middle Miocene deepwater paleoceanography in the southwest Pacific: Relations with east Antarctic ice sheet development. *Paleoceanography* **10**, 1095–1113 (1995).
25. Flower, B. P. & Kennett, J. P. Middle Miocene ocean/climate transition: High-resolution oxygen and carbon isotopic records from DSDP Site 588A, Southwest Pacific. *Paleoceanography* **8**, 811–843 (1993).
26. Shackleton, N. J. *Marine Petroleum Source Rocks* (eds Brooks, J. et al.) 423–434 (Spec. Publ. 26, Geol. Soc., London, 1987).

Supplementary Information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from Mary Sheehan at the London editorial office of Nature.

Acknowledgements. We thank S. Clemens, P. DeMenocal, P. Koch, N. Shackleton, J. Revenaugh and Q. Williams for comments and discussion; K. Vencil for technical assistance; and S. D'Hondt, L. Hinnov and K. Miller for reviews that significantly improved the manuscript. Samples for this project were provided by the Ocean Drilling Program. This work was supported by the USSSP and NSF.

Correspondence and requests for data should be addressed to J.C.Z. (e-mail: jzachs@earthsci.ucsc.edu).

Geodynamic estimates of the viscosity of the Earth's inner core

Bruce A. Buffett

Department of Earth and Ocean Sciences, University of British Columbia, Vancouver, BC, Canada, V6T 1Z4

Recent seismological studies^{1,2} have suggested that the inner core is rotating relative to the bulk of the Earth, a situation which (according to numerical simulations³) may be sustained by convective flow in the liquid outer core. On the other hand, large gravitational forces due to the heterogeneous distribution of mass in the Earth's mantle should be sufficient to keep the inner core aligned with the mantle⁴. Here I show that the differential rotation of the inner core can be reconciled with these strong gravitational forces by allowing the shape of the inner core to adjust as it rotates, so permitting an estimate of the effective viscosity of this innermost region of the Earth. The inferred rotation rate constrains the viscosity of the inner core to be less than 10^{16} Pa s or greater than 10^{20} Pa s, as two different dynamical regimes are possible. The viscosity estimates for these two regimes have very different implications for the origin of seismic anisotropy in the inner core.

Investigations of the inner core reveal that the elastic velocity in the polar direction is faster than that in the equatorial direction by $\sim 3\%$ (refs 5–7). This anisotropy has a cylindrical symmetry about an axis which is slightly misaligned with the Earth's rotation axis⁷. Two recent studies^{1,2} have detected systematic variations in the travel time of seismic waves that pass through the inner core. These variations have been attributed to a differential rotation of the inner core which carries with it the inclined symmetry axis of the anisotropy. The inferred rotation rate of the inner core, relative to the mantle, is roughly 1° per year.

The search for systematic changes in seismic travel times was motivated, in part, by a three-dimensional, numerical simulation of the geodynamo⁸. These calculations predict an eastward rotation of the inner core with an angular velocity of 10^{-9} rad s⁻¹, consistent with the values subsequently reported by the seismological studies. The predicted rotation is maintained by magnetic coupling between the inner core and an eastward flow of the overlying liquid outer core^{3,9}.

A complication is introduced to our present understanding of inner-core rotation by the presence of mass anomalies in the mantle. These mass anomalies cause gravitational perturbations that distort surfaces of constant potential throughout the Earth. Mass ana-

lies derived from seismic tomography have been used successfully in viscous models of the mantle to reproduce the low-degree components of the equipotential surface at mean sea level. The same calculations can yield a 100-m radial variation in the equipotential surface at the top of the inner core¹⁰. To a good approximation, the inner core is in thermodynamic and hydrostatic equilibrium with the surrounding fluid when conditions are averaged over dynamic fluctuations. In this case the time-averaged solidification and equipotential surfaces should coincide. The resulting (hydrostatic) structure on the inner-core boundary can produce surprisingly large gravitational forces when the inner core is rotated out to its equilibrium alignment with the mantle. The restoring torque Γ_g on the inner core can be as large as 10^{21} N m (ref. 4), whereas the dynamo process in the outer core is likely to produce an electromagnetic torque Γ_b on the inner core of only 10^{19} N m, based on recent geodynamo calculations³. The torque Γ_b would rotate the inner core out of alignment with the mantle, but a balancing gravitational torque is achieved with an angular deviation $\Delta\phi$ of only 0.01 rad. Thus the inner core should remain gravitationally locked to the mantle.

The seismically inferred rotation of the inner core can be accommodated by allowing the inner-core shape to adjust as it rotates through the gravitational potential imposed by the overlying mantle. The shape of the inner core is expressed as a function of co-latitude θ and longitude ϕ by boundary topography $h(\theta, \phi)$ on a rotationally symmetric ellipsoid. It is most convenient to express $h(\theta, \phi)$ in a coordinate frame which is fixed to the mantle. If the inner core is stationary relative to the mantle, then h slowly adjusts to the equilibrium configuration h^e . Subsequent changes in h are caused by rotation of the inner core relative to the mantle and by relaxation of the inner-core shape when h deviates from h^e . When the boundary topography is described in terms of spherical harmonics¹¹ $Y_l^m(\theta, \phi)$, according to $h = h_l Y_l^m(\theta, \phi)$, we can write the rate of change of the coefficient h_l as

$$\frac{\partial h_l}{\partial t} = -\frac{(h_l - h_l^e)}{\tau} - i m \phi h_l \quad (1)$$

where ϕ is the angular velocity of the inner core, $i = \sqrt{-1}$ and τ is the characteristic time for h_l to adjust to the equilibrium value h_l^e . This adjustment may occur either by melting and solidification at the inner-core boundary or by viscous relaxation. So long as τ is constant it is not necessary to distinguish between these two possibilities in order to solve equation (1).

Two limiting cases of the steady solution

$$h_l = h_l^e (1 + i m \phi \tau)^{-1} \quad (2)$$

permit the inner core to rotate relative to the mantle. The first occurs when $\phi \tau \ll 1$. In this case h_l can adjust rapidly to remain nearly aligned with h_l^e , while mass elements that are imprinted with seismic anisotropy can rotate at the rate ϕ relative to the mantle. When h and h^e are misaligned by a small angle $\Delta\phi$, we can

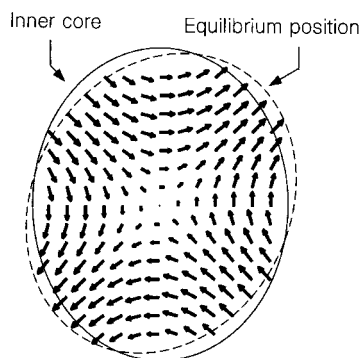


Figure 1 Velocity field induced by a misalignment of the inner core from its equilibrium position. Arrows denote the direction of flow and their length indicates the relative amplitude. The combined effect of rotation and relaxation allows material elements in the inner core to drift relative to the mantle. The total strain induced by the viscous flow is small because the shape of the inner core deviates from rotational symmetry by only $\Delta R/R \approx 10^{-4}$. Heating associated with the rate of strain is 5×10^9 W when the viscosity is 3×10^{16} Pa s.

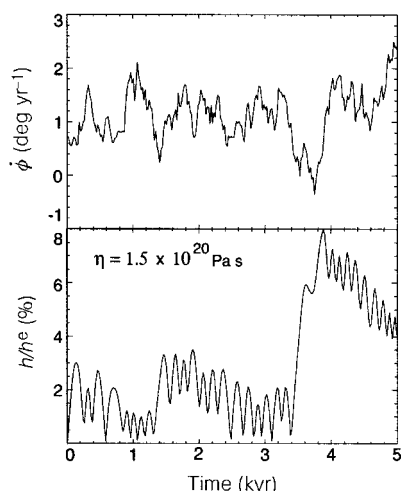


Table 1 Physical properties of the inner core

Parameter	Symbol	Value
Radius	R	1.2×10^6 m
Boundary topography	ΔR	100 m
Latent heat	L	10^6 J kg $^{-1}$
Density	ρ	1.2×10^4 kg m $^{-3}$
Density jump at R	$\Delta \rho$	600 kg m $^{-3}$
Gravity at R	g	4.4 m s $^{-2}$

also represent the topography by $h_i \approx h_i^e e^{-im\Delta\phi}$ and obtain from equation (2)

$$\Delta\phi = \phi\tau \quad (3)$$

which is valid to first-order in $\phi\tau$. As ϕ is ~ 0.017 rad year $^{-1}$ and $\Delta\phi$ should not exceed 0.01 rad, we obtain a relaxation time τ of ~ 0.6 years or less.

Such a short relaxation time is difficult to accommodate by melting and solidification at the inner-core boundary because of the enormous amount of latent heat that must be redistributed around the surface of the inner core. This heat may be transported by radial convective motions or redistributed by tangential flows due to lateral temperature variations over the surface of the inner core. In either case, the transport of latent heat demands a net heat flux $q = \rho L \Delta R / \tau$, where L is the latent heat, ρ is the density and ΔR is the topographic change accomplished in time τ . Using the parameter values listed in Table 1, we require a heat flux of 6×10^4 W m $^{-2}$ to achieve a relaxation time of 0.6 years. Such a heat flux is far too high to be driven by typical temperature variations of only a few millikelvin.

Viscous relaxation provides a more plausible mechanism for allowing changes in the shape of the inner core. The process is analogous to isostatic adjustments of the mantle, for which a well known theoretical framework is established. It is assumed that the inner core can be modelled as a self-gravitating, incompressible viscous sphere that is slightly distorted by a perturbing gravitational potential from the overlying mantle. Numerical calculations, based on the formulation of Hager and Clayton¹², are used to predict the flow induced by small angular misalignments. The velocity field in an isoviscous inner core is shown in Fig. 1 for the largest component of boundary topography ($l = 2$, $m = 2$). The estimate of τ for a uniform inner core with constant viscosity η is given by

$$\tau = \frac{C_1 \eta}{\Delta \rho g R} \quad (4)$$

where $\Delta \rho$ and g are the density jump and local gravity at the radius R of the inner-core boundary, and $C_1 = 1.9$ is a constant determined

Figure 2 Adjustment of the inner core when the rotation rate ϕ is variable. Variations in rotation rate (ϕ ; top panel) are modelled as a first-order Gauss–Markov process with a correlation time of 500 years. Excursions from the mean rotation rate of 1° yr $^{-1}$ are $\pm 1^\circ$ yr $^{-1}$. The resulting change in the amplitude of the boundary topography h (normalized by the equilibrium value, h^e ; bottom panel) is determined by numerically integrating equation (1) with τ determined for the case of a uniform viscosity $\eta = 1.5 \times 10^{20}$ Pa s. During intervals of slow rotation the boundary topography grows, increasing the strength of the gravitational torque. Using the parameter values adopted in this study, the gravitational torque is $8\times$ larger than the electromagnetic torque when $h/h^e \approx 8\%$.

from the numerical calculations. A relaxation time of 0.6 years or less implies a viscosity of less than 3×10^{16} Pa s because all other parameters in equation (4) are relatively well constrained.

Lower estimates of viscosity are required if the relaxation is confined to a layer at the top of the inner core. Such a layer might coincide with a two-phase mushy region¹³ or a region of recently crystallized, fine-grained solid¹⁴. When the thickness d of the layer is small compared with R , the relaxation time may be approximated by

$$\tau = \frac{C_2 \eta}{\Delta \rho g R} \left(\frac{R}{d} \right)^3 \quad (5)$$

where $C_2 = 0.21$ is a numerical constant. For example, a 100-km layer must have a viscosity less than 2×10^{14} Pa s to accommodate the observed rotation of the inner core.

A very different view of inner-core rotation is inferred from equation (2) in the limit $\phi\tau \gg 1$. In this case the relatively rapid rotation ensures that h is much smaller than h^e , so that gravitational locking becomes much weaker. The gravitational torque becomes comparable to the electromagnetic torque when h is reduced from the equilibrium value by a factor of 100. To achieve $h/h^e < 0.01$ in equation (2), we require a relaxation time in excess of 2,800 years for $m = 2$, corresponding to a viscosity of 1.5×10^{20} Pa s or more. Such a high viscosity poses problems if the inner core does not rotate uniformly. During intervals of slower rotation the topography can grow, thereby locking the inner core to the mantle. Once the highly viscous inner core becomes locked it is likely to remain so. Figure 2 demonstrates this possibility by letting ϕ vary in equation (1) as a first-order Gauss–Markov process about a mean value of 1° per year. A numerical integration for h_2 reveals that dramatic increases in topography are possible. Given the vulnerability of the high-viscosity inner core to locking, the low-viscosity estimate is favoured.

How does this geodynamic estimate of viscosity compare with other inferences? The most probable mechanism for creep at high temperature and low stress is the diffusion of lattice vacancies¹⁵. Temperature and pressure both influence the effective viscosity, and their role is often included in rheological models by referring the temperature to the melting temperature T_m ; the pressure dependence enters implicitly through its influence on T_m . The ratio T/T_m probably exceeds 0.9 throughout the inner core, so high-temperature experiments on iron at ambient pressure should provide one basis for comparison. Typical experimental estimates¹⁶ yield a viscosity of 10^{16} Pa s for a grain size of roughly 5 mm. This estimate of grain size is plausible, but uncertain because of the pressure extrapolation.

Substantially larger grain sizes have recently been advocated¹⁷, but such physical conditions are incompatible with the low-viscosity

estimate because larger grain sizes are increasingly resistant to deformation. Although it is possible for large crystals to grow in the deepest part of the inner core, this would confine the viscous flow to the outermost region. Low geodynamic bounds on the viscosity of a layer imply much smaller grain sizes in this region. For example, a grain size of 0.5 mm or less is required, if the flow is confined to a 100-km-deep layer.

A complementary estimate of viscosity is deduced by assuming that the mechanism responsible for viscous creep is also responsible for shear attenuation of seismic waves^{14,18}. If the inner core behaves like a Maxwell solid, then we find that a quality factor Q_μ of roughly 100 for low-frequency free oscillations¹⁹ implies a viscosity of 10^{15} Pa s. This value is a lower bound because all of the seismic attenuation is attributed to the creep mechanism. By contrast, the low-viscosity estimate of 3×10^{16} Pa s can be viewed as an upper bound for the preferred mode of adjustment. These two estimates are mutually consistent, but the common range of viscosity is remarkably narrow.

Geodynamic estimates of inner-core viscosity are important for several purposes. Dynamical processes such as thermal convection²⁰ and viscous compaction²¹ are strongly influenced by the value of viscosity. Moreover, the small grain sizes implied by such estimates yield insights into crystal growth in the inner core. Restrictions on the growth of iron grains may point to the importance of deformational stresses or chemical impurities. Small grain sizes and low viscosities also have important consequences for the source of seismic anisotropy in the inner core. Persistent deformation of the inner core owing to its rotation relative to the mantle does not favour the high degree of crystal alignment that is needed in the high-pressure (hexagonal close-packed) phase of iron to explain the anisotropy¹⁷. A search for alternative explanations of the anisotropy is thus required. Unravelling the clues left behind by the growth of the inner core offers the hope of new insights into the evolution of the Earth's deep interior. □

Received 31 January; accepted 1 July 1997.

1. Song, X. & Richards, P. Seismological evidence for the rotation of the Earth's inner core. *Nature* **382**, 221–224 (1996).
2. Su, W.-J., Dziewonski, A. M. & Jeanloz, R. Planet within a planet: rotation of the inner core of the Earth. *Science* **274**, 1883–1887 (1996).
3. Glatzmaier, G. & Roberts, P. H. Rotation and magnetism of Earth's inner core. *Science* **274**, 1887–1891 (1996).
4. Buffett, B. A. A mechanism for decade fluctuations in the length of day. *Geophys. Res. Lett.* **23**, 3803–3806 (1996).
5. Creager, K. C. Anisotropy of the inner core from differential travel times of phases PKP and PKIKP. *Nature* **356**, 309–314 (1992).
6. Song, X. & Helmberger, D. V. Depth dependence of anisotropy of Earth's inner core. *J. Geophys. Res.* **100**, 9805–9816 (1995).
7. Su, W.-J. & Dziewonski, A. M. Inner core anisotropy in three dimensions. *J. Geophys. Res.* **100**, 9831–9852 (1995).
8. Glatzmaier, G. & Roberts, P. H. An anelastic evolutionary geodynamo simulation driven by compositional and thermal convection. *Physica D* **97**, 81–94 (1996).
9. Aurnou, J. M., Brito, M. & Olson, P. L. Mechanics of inner core super-rotation. *Geophys. Res. Lett.* **23**, 3401–3404 (1996).
10. Forte, A. & Peltier, W. R. Viscous flow models of global geophysical observables, 1 Forward problems. *J. Geophys. Res.* **99**, 21857–21877 (1991).
11. Edmonds, A. R. *Angular Momentum in Quantum Mechanics* (Princeton Univ. Press, 1960).
12. Hager, B. H. & Clayton, R. W. in *Mantle Convection* (Gordon & Breach, New York, 1989).
13. Loper, D. E. & Roberts, P. H. A study of conditions at the inner core boundary of the Earth. *Phys. Earth Planet. Inter.* **24**, 302–307 (1981).
14. Yoshida, S., Sumita, I. & Kumazawa, M. Growth model of the inner core coupled with outer core dynamics and the resulting elastic anisotropy. *J. Geophys. Res.* **101**, 2805–28103 (1996).
15. Poirier, J.-P. *Creep of Crystals: High-temperature Deformation Processes in Metals, Ceramics, and Minerals* (Cambridge Univ. Press, 1985).
16. Frost, H. J. & Ashby, M. F. *Deformation Mechanism Maps* Ch. 4 (Pergamon, Oxford, 1983).
17. Stixrude, L. & Cohen, R. E. High-pressure elasticity of iron and anisotropy of Earth's inner core. *Science* **267**, 1972–1975 (1995).
18. Anderson, D. L. Earth's viscosity. *Science* **151**, 321–322 (1966).
19. Widmer, R., Masters, G. & Gilbert, F. Spherically symmetric attenuation within the Earth from normal mode data. *Geophys. J. Int.* **104**, 541–553 (1991).
20. Jeanloz, R. & Wenk, H.-R. Convection and anisotropy of the inner core. *Geophys. Res. Lett.* **15**, 72–75 (1988).
21. Sumita, I., Yoshida, S., Kumazawa, M. & Hamano, Y. A model for sedimentary compaction of a viscous medium and its application to inner-core growth. *Geophys. J. Int.* **124**, 502–524 (1996).

Acknowledgements. I thank J. Bloxham, B. Hager, P. Richards and D. Stevenson for comments and suggestions. J. Bloxham and D. Stevenson independently suggested the possibility of the high-viscosity solution. This work was supported by NSERC.

Correspondence and requests for materials should be addressed to the author (e-mail: buffett@geop.ubc.ca).

Deep marine biosphere fuelled by increasing organic matter availability during burial and heating

Peter Wellsbury, Kim Goodman, Tanja Barth*, Barry A. Cragg, Stephen P. Barnes & R. John Parkes

Department of Geology, University of Bristol, Wills Memorial Building, Queens Road, Bristol BS8 1RJ, UK

* Department of Chemistry, University of Bergen, Allg. 41, N-5007 Bergen, Norway

Deep-sea sediments become apparently more hostile to life with increasing depth as temperature and pressure rise, and organic matter becomes increasingly recalcitrant. Demonstrations of high bacterial populations in deep sediments^{1,2} may thus appear enigmatic. How, then, can the continued presence of active bacterial populations in deep sediments that are over 10 million years old be explained? Although volatile fatty acids, particularly acetate, are important intermediates in the anaerobic degradation of organic matter^{3,4}, their concentrations are kept very low in sediments ($<15 \mu\text{M}$) by rapid bacterial consumption^{5,6}. Here we show that heating surface coastal marine sediments to simulate increasing temperature during burial produces an increase of over three orders of magnitude in acetate concentration and increases bacterial activity. We found that pore-water acetate concentration at two sites in the Atlantic Ocean increased at depths below about 150 m and was associated with a significant stimulation in bacterial activity. Comparing these acetate concentrations to *in situ* temperatures confirmed that there was a notable generation of acetate associated with temperature increases during burial. This was supported by heating experiments with deep sediments. Thus, acetate generation from organic matter during burial may explain the presence of a deep bacterial biosphere in marine sediments, and could underpin an even deeper and hotter biosphere than has previously been considered.

In anaerobic marine sediments, volatile fatty acids (VFA) can reflect the intensity and route of organic matter degradation, as they are important metabolic intermediates in anaerobic bacterial metabolism^{3,4}. The most important is acetate, which accounts for $>60\%$ of sulphate reduction⁴, the dominant anaerobic process in marine sediments⁵. Closely coupled production and consumption of acetate in near-surface sediments keeps pore-water concentrations very low ($<15 \mu\text{M}$; ref. 6); in deeper sediments, even lower acetate concentrations would be expected. This is reinforced by the observation that bacterial methane production, which occurs deeper than sulphate reduction in marine sediments, is predominantly from H_2 and CO_2 rather than acetate⁷. However, VFA are known to exist at high concentrations in deep oil-reservoir-formation waters, considered to be due to abiological organic matter cracking at temperatures above 80°C ^{8,9}.

In heating experiments (0 – 100°C) with surface coastal marine sediments, we investigated the possibility of acetate being generated from organic matter as a result of the increasing thermal gradient during sediment burial (Fig. 1), and within the temperature range for bacterial activity. After only 7 days, notable amounts of acetate ($24,000 \mu\text{M}$) were produced between 10 and 60°C (temperature range within that for meso- and thermophilic bacteria) compared to the *in situ* concentration of $12 \mu\text{M}$. This production is considered to be bacterial, as little acetate was produced at temperatures $>60^\circ\text{C}$, and acetate concentrations reflected a typical bacterial growth curve with increasing temperature. This acetate was simultaneously being

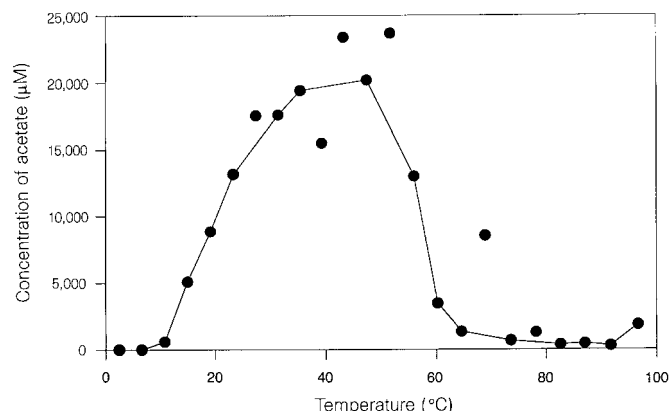


Figure 1 Generation of pore-water acetate from a coastal near-surface sediment exposed to low-temperature heating. Material taken from the uppermost 1 cm of an intertidal mudflat at Southdown, Tamar Estuary, UK, was slurried (25% v/v) with sterile anoxic artificial seawater and incubated in a thermal gradient system²⁸ for 7 days. Pore-water acetate concentrations were determined by enzymatic high-pressure liquid chromatography (HPLC)¹¹. Active sulphate reduction occurred during incubation; as acetate is the principal substrate for this activity²¹ in the Tamar Estuary, acetate concentrations were compensated for acetate consumption based on the amount of sulphate removed, and dilution due to slurring.

used as an energy source by sulphate-reducing bacteria (92% of acetate production was consumed).

As a consequence of these results, concentrations of pore-water acetate were determined in deeper sediments than had been previously analysed¹⁰ to investigate the effect of temperature increases during burial (Blake Ridge, Atlantic Ocean, ODP Leg 164, Site 995; Fig. 2). Although near-surface concentrations were within the expected low range (7 μM at 0.5 m) and remained close to these levels to 82.5 m, acetate concentrations began to increase (68 μM at 154 m) in deeper sediment and continued to increase with increasing depth. Concentrations reached 15,285 μM at 691 m, an increase of over three orders of magnitude (2,248 times). These measurements were confirmed by two independent techniques; ion chromatography and a specific enzymatic technique. The enzymatic technique determines acetate in a biologically available form^{6,11}, and indeed some of this acetate is being metabolized in the deep sediments. For example, methanogenesis from acetate below about 400 m was an order of magnitude greater than the maximum surface rate (1,240 nmol ml⁻¹ day⁻¹ at 599 m, 183 nmol ml⁻¹ day⁻¹ at 1.65 m, respectively). Acetate oxidation to CO₂ was stimulated to an even greater extent (Fig. 2). Below about 400 m, increases in other VFA were also detected.

Bacterial processes are complicated in these deeper sediments by the presence of methane hydrate between 445 m (base of bottom-simulating reflector) and 190 m (top of hydrate stability field¹²). Previous research has demonstrated that deep bacterial populations are stimulated within methane hydrate zones² and this is also evident at Blake Ridge (Fig. 2). Generally, a range of bacterial activities (methanogenesis from carbon dioxide¹³, methanogenesis from acetate⁶, methane oxidation¹³) and bacterial growth (thymidine incorporation¹⁴) are stimulated around the hydrate (370 to 500 m) and this results in a significant increase in bacterial populations (to 5.6×10^7 cells per ml). In deeper layers, where acetate increases further, methane production from acetate, acetate mineralization and methane oxidation continue to increase. Methane production from acetate, however, is much greater than methane

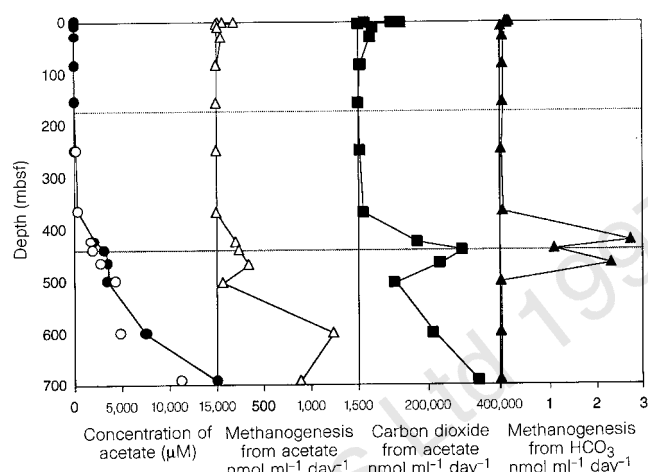


Figure 2 Depth profiles of pore-water acetate concentrations and rates of bacterial activity in sediments from Blake Ridge, ODP Leg 164, Site 995. Acetate determinations were performed on filter-sterilized (0.1 μm) pore-water samples by enzymatic HPLC¹¹ (filled circles) and ion-exclusion chromatography²⁹ (empty circles). Rates of methane production from bicarbonate (filled triangles) and acetate (empty triangles), and acetate oxidation to CO₂ (squares) were determined using H¹⁴CO₃⁻ and [1-(2)¹⁴C]acetate, respectively, as radiotracers, injected into subcores of sediment and incubated for a range of time periods before determination of ¹⁴CH₄ and ¹⁴CO₂^{6,13}. Dotted lines indicate the upper and lower (bottom-simulating reflector) boundaries of the gas hydrate stability field.

oxidation, which is consistent with increased methane concentrations at depth¹².

When plotted against temperature¹⁵, rather than depth, Site 995 pore-water acetate concentrations show a remarkable similarity to the acetate generated by heating coastal marine sediments (Fig. 3). In addition, the pore-water acetate depth profile (determined by isotachophoresis¹⁶) from another site (997) on the Blake Ridge was similar to that of Site 995. Any disparity was probably the result of differences in the thermal gradient at the sites (32 and 37 °C km⁻¹ at 995 and 997, respectively¹⁵) because, when plotted against temperature, the acetate concentrations at the two sites are identical (not significantly different at $P > 0.01$). This demonstrates that deep-sourced acetate may be a general feature of marine sediments and that generation is closely coupled to increased temperature during burial. Heating even more deeply buried sediments to much higher temperatures (such as Kimmeridge Clay at 250–350 °C) releases further acetate (5.7–18.9 mmol kg⁻¹ rock⁸) through abiological mechanisms, providing a potential for continued acetate generation throughout sediment burial. As bacteria can grow up to about 113 °C, and possibly even higher^{17,18}, acetate generation from organic matter potentially provides an energy source for a deep hot biosphere¹⁹.

If this theory of low-temperature acetate generation during organic matter burial is correct, then one would expect that deeper sediments, already subjected to increased temperatures and some acetate production, would generate less acetate on further heating compared to a near-surface sediment. To test this, we heated sediments from different depths (0.13, 365, 690 m) from the same site (995; Blake Ridge), for which the pore-water profiles were obtained, together with near-surface samples from a Pacific Ocean ODP site (Leg 146, Cascadia Margin). After only 7 days incubation of Blake Ridge sediments, considerable acetate was generated, and after 21 days (Fig. 4) the amounts of acetate produced were similar to the concentrations present in deep sediment pore-water (Fig. 2), although, unsurprisingly, somewhat lower than those produced on heating coastal sediments, which contain highly labile organic

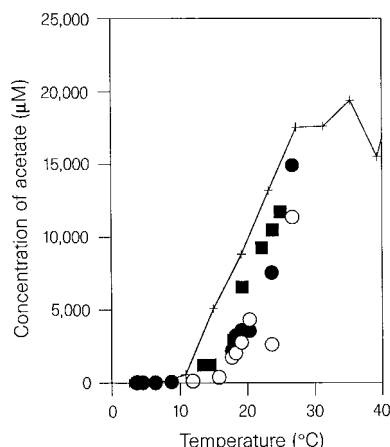


Figure 3 Comparison of *in situ* pore-water acetate concentrations from Blake Ridge sediments with experimental low-temperature acetate generation from near-surface coastal sediments. Site 995 was determined by enzymatic HPLC¹¹ (filled circles) and ion-exclusion chromatography²⁹ (empty circles), and Site 997 (squares) by isotachophoresis¹⁶. For comparison, the solid line is taken from the surface sediment heating experiments shown in Fig. 1.

matter (Fig. 1). Acetate generation decreased, as expected, with increasing depth. In addition, generation from the deepest sample (690 m) required a much higher temperature (55 °C) compared to the near-surface sample (0.13 m), where maximal acetate generation occurred at 30 °C. Generation at 365 m was intermediate between these two. Acetate production also occurred in heated near-surface Pacific Ocean sediments from the Cascadia Margin (acetate concentration increased from 13 μM to 1,087 μM after 22 days at 30 °C).

Methane production from acetate in deep sediments should be reflected in the isotope composition of methane gas, and $\delta^{13}\text{C}$ values confirm CH_4 in the Blake Ridge to be biogenic²⁰. However, based on a combination of $\delta^{13}\text{C}$ and δD values, the bulk of biogenic methane in marine sediments has been considered to originate from $\text{H}_2:\text{CO}_2$ rather than acetate⁷. Other sources of CH_4 in marine sediments have been ascribed to abiological thermogenic processes; such methane, although it has a similar $\delta^{13}\text{C}$ value to acetate methanogenesis (−50‰), may be discriminated on the basis of its δD (thermogenic < −250‰, acetate > −250‰; ref. 7). Recent evidence²¹ has shown, however, that there is significant deuterium exchange during acetate methanogenesis. Hence, δD values may not always be an effective means of distinguishing between abiological thermogenic and biological acetoclastic methanogenesis. Therefore, it is possible that acetoclastic CH_4 generation in deep marine sediments has previously been underestimated.

Our results demonstrate that relatively low-temperature heating during burial increases the bioavailability of sedimentary organic matter and this is reflected in marked increases in pore-water acetate concentrations. This process would continue during burial and, together with acetate consumption, provides a bacterial energy source that continues to increase with depth and would explain the presence of the significant bacterial populations now known to be widespread in deep marine sediments^{1,2,18,22–26}. These results also offer the prospect of viable bacterial populations being present in marine sediments much deeper than the deepest samples currently analysed (750 m; ref. 27). As temperature increases in even deeper sediments, high-temperature bacteria may develop (thermophiles and hyperthermophiles, 45 to 113 °C¹⁷). These prokaryotes could continue to exploit buried organic matter made more bioavailable owing to heating, even into the oil window (100 to 150 °C). Hence,

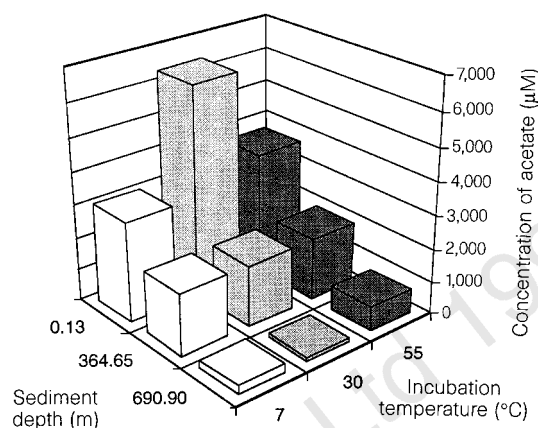


Figure 4 Generation of acetate from deep-sediment samples subjected to low-temperature heating. Sediment from ODP Leg 164, site 995 (Blake Ridge, 0.13, 365 and 690 m depth), was slurried with sterile, anoxic mineral salts solution and incubated for 22 days before pore-water extraction by centrifugation under nitrogen and determination of pore-water acetate concentrations by enzymatic HPLC¹¹. To enable comparison with *in situ* profiles, acetate concentrations were adjusted for dilution during slurring.

deep sediment bacteria may play a role in organic matter maturation, a prelude to oil and gas generation. □

Received 19 May; accepted 19 June 1997.

- Parkes, R. J. *et al.* Deep bacterial biosphere in Pacific Ocean sediments. *Nature* **371**, 410–413 (1994).
- Cragg, B. A. *et al.* Bacterial populations and processes in sediments containing gas hydrates (ODP Leg 146: Cascadia Margin). *Earth Planet. Sci. Lett.* **139**, 497–507 (1996).
- Sørensen, J., Christensen, D. & Jørgensen, B. B. Volatile fatty acids and hydrogen as substrates for sulfate-reducing bacteria in anaerobic marine sediment. *Appl. Environ. Microbiol.* **42**, 5–11 (1981).
- Winfrey, M. R. & Ward, D. M. Substrates for sulfate reduction and methane production in intertidal sediments. *Appl. Environ. Microbiol.* **45**, 193–199 (1983).
- Jørgensen, B. B. Mineralisation of organic matter in the sea bed: The role of sulphate reduction. *Nature* **296**, 643–645 (1982).
- Wellsbury, P. & Parkes, R. J. Acetate bioavailability and turnover in an estuarine sediment. *FEMS Microbiol. Ecol.* **17**, 85–94 (1995).
- Whiticar, M. J., Faber, E. & Schoell, M. Biogenic methane formation in marine and freshwater environments: CO_2 reduction vs. acetate fermentation: Isotope evidence. *Geochim. Cosmochim. Acta* **50**, 693–709 (1986).
- Coolles, G. P., Mackenzie, A. S. & Parkes, R. J. Non-hydrocarbons of significance in petroleum exploration: Volatile fatty acids and non-hydrocarbon gases. *Mineral. Mag.* **51**, 483–493 (1987).
- Borgund, A. E. & Barth, T. Generation of short-chain organic-acids from crude-oil by hydrous pyrolysis. *Org. Geochem.* **21**, 943–952 (1994).
- Whelan, J. K. *et al.* Evidence for sulfate-reducing and methane-producing organisms in sediments from Sites 618, 619 and 622. *Init. Rep. Deep-Sea Drill. Proj.* **96**, 767–775 (1986).
- King, G. M. Measurement of acetate concentrations in marine pore water by using an enzymatic approach. *Appl. Environ. Microbiol.* **57**, 3476–3481 (1991).
- Dickens, G. R. *et al.* Direct measurements of *in situ* methane quantities in a large gas-hydrate reservoir. *Nature* **385**, 426–428 (1997).
- Cragg, B. A. *et al.* Bacterial biomass and activity profiles within deep sediment layers. *Proc. ODP Sci. Res.* **112**, 607–619 (1990).
- Wellsbury, P., Herbert, R. A. & Parkes, R. J. Bacterial activity and production in near-surface estuarine and freshwater sediments. *FEMS Microbiol. Ecol.* **19**, 203–214 (1996).
- Paull, C. K. *et al.* Leg 164—gas hydrate sampling on the Blake Ridge and Carolina Rise. *Proc. ODP Init. Rep.* **164**, 623 (1996).
- Barth, T. Quantitative determination of volatile carboxylic acids in formation water by isotachophoresis. *Anal. Chem.* **59**, 2232–2237 (1987).
- Stetter, K. O. Hyperthermophiles in the history of life. *Ciba Found. Symp.* **202**, 1–18 (1996).
- Cragg, B. A. & Parkes, R. J. Bacterial profiles in hydrothermally active deep sediment layers from Middle Valley (N.E. Pacific) Sites 857 and 858. *Proc. ODP Sci. Res.* **139**, 509–516 (1993).
- Haridon, S. L., Reysenbach, A., Glat, P., Prieur, D. & Jeanthon, C. Hot subterranean biosphere in continental oil reservoir. *Nature* **377**, 223–224 (1995).
- Galimov, E. M. & Kvendovden, K. A. Concentrations and carbon isotopic compositions of CH_4 and CO_2 in gas from sediments of the Blake Outer Ridge, Deep-Sea Drilling Project Leg 76. *Init. Rep. Deep-Sea Drill. Proj.* **76**, 403–407 (1983).
- de Graaf, W., Wellsbury, P., Parkes, R. J. & Capenberg, T. E. Comparison of acetate turnover in methanogenic and sulfate-reducing sediments by radio- and stable-isotope-labeling and specific inhibitors: Evidence for isotopic exchange. *Appl. Environ. Microbiol.* **62**, 772–777 (1996).
- Cragg, B. A. *et al.* Bacterial profiles in deep sediments of the Santa Barbara Basin Site 893. *Proc. ODP Sci. Res.* **146**, 139–144 (1995).
- Cragg, B. A. *et al.* The impact of fluid and gas venting on bacterial populations and processes in sediments from the Cascadia Margin accretionary system (sites 888–892) and the geochemical consequences. *Proc. ODP Sci. Res.* **146**, 399–411 (1995).
- Furnes, H., Thorseth, I. H., Tørrvik, O., Torsvik, T. & Fisk, M. R. Microbial activity in the alteration of glass from pillow lavas from Hole 896A. *Proc. ODP Sci. Res.* **148**, 191–214 (1996).

25. Giovannoni, S. J., Fisk, M. R., Mullins, T. D. & Furnes, H. Genetic evidence for endolithic microbial life colonising basaltic glass/seawater interfaces. *Proc. ODP Sci. Res.* **148**, 207–214 (1996).
26. Rochelle, P. A., Fry, J. C., Parkes, R. J. & Weightman, A. J. DNA extraction for 16S rRNA gene analysis for determining genetic diversity in deep sediment communities. *FEMS Microbiol. Lett.* **100**, 59–66 (1992).
27. Parkes, R. J. *et al.* Biogeochemical processes in gas hydrate zones. *Abstracts, Exploration and exploitation of non-living deep-sea marine resources*. Challenger Society for Marine Science & Geological Society, Marine Studies Group (Geological Society, London, 1996).
28. Jørgensen, B. B., Isaksen, M. F. & Jannasch, H. W. Bacterial sulfate reduction above 100 °C in deep-sea hydrothermal sediments. *Science* **258**, 1756–1757 (1992).
29. Parkes, R. J. & Taylor, J. Analysis of volatile fatty acids by ion-exclusion chromatography, with special reference to marine pore water. *Mar. Biol.* **77**, 113–118 (1983).

Acknowledgements. We thank ODP for allowing us to obtain samples on Leg 164; T. Woodward and I. Mather for their assistance with sample collection and initial handling in the laboratory; J. Maxwell, E. Shock and J. Whelan for comments on the original manuscript; P. Egeberg for facilitating analysis of Site 997; M. Isaksen for details of the thermal gradient system, and F. Wheeler for its construction. This work was funded by the European Union Environment Programme and the Natural Environment Research Council (UK).

Correspondence and requests for materials should be addressed to R.J.P. (e-mail: j.parkes@bris.ac.uk).

The fate of carbon in grasslands under carbon dioxide enrichment

Bruce A. Hungate^{*,†}, Elisabeth A. Holland[‡], Robert B. Jackson[§], F. Stuart Chapin III[¶], Harold A. Mooney^{||} & Christopher B. Field[†]

^{*} Department of Integrative Biology, University of California, Berkeley, California 94720, USA

[‡] Atmospheric Chemistry Division, National Center for Atmospheric Research, Boulder, Colorado 80307, USA

^{||} Department of Biological Sciences, Stanford University, Stanford, California 94305, USA

[§] Department of Botany, University of Texas at Austin, Austin, Texas 78713, USA

[¶] Carnegie Institution of Washington, Department of Plant Biology, Stanford, California 94305, USA

[†] Present address: Smithsonian Environmental Research Center, Edgewater, Maryland 21037, USA.

The concentration of carbon dioxide (CO₂) in the Earth's atmosphere is rising rapidly¹, with the potential to alter many ecosystem processes. Elevated CO₂ often stimulates photosynthesis², creating the possibility that the terrestrial biosphere will sequester carbon in response to rising atmospheric CO₂ concentration, partly offsetting emissions from fossil-fuel combustion, cement manufacture, and deforestation^{3,4}. However, the responses of intact ecosystems to elevated CO₂ concentration, particularly

the below-ground responses, are not well understood. Here we present an annual budget focusing on below-ground carbon cycling for two grassland ecosystems exposed to elevated CO₂ concentrations. Three years of experimental CO₂ doubling increased ecosystem carbon uptake, but greatly increased carbon partitioning to rapidly cycling carbon pools below ground. This provides an explanation for the imbalance observed in numerous CO₂ experiments, where the carbon increment from increased photosynthesis is greater than the increments in ecosystem carbon stocks. The shift in ecosystem carbon partitioning suggests that elevated CO₂ concentration causes a greater increase in carbon cycling than in carbon storage in grasslands.

In most ecosystems, leaf and canopy gas-exchange measurements indicate that there is increased carbon uptake in response to experimental doubling of CO₂ concentration^{5,6}. However, increased carbon uptake in these one- to ten-year experiments does not necessarily indicate a large potential for carbon storage over 50–100 years, the predicted timescale of atmospheric CO₂ doubling⁷. Ecosystem sequestration of carbon over 50–100 years requires delivery of the extra carbon to large pools with slow turnover (wood and soil organic matter)⁸. However, net gas-exchange measurements do not reveal whether the extra carbon fixed in response to elevated CO₂ is distributed to these pools. The partitioning of this extra carbon among pools with varying turnover time is a critical controller of the potential for terrestrial ecosystems to increase long-term carbon storage.

A consequence of CO₂-stimulated plant growth may be an increased demand for below-ground resources (water and mineral nutrients). If the demands for below-ground resources are not met by increased resource availability or efficiency of resource use^{9,10}, or if growth potential is constrained¹¹, plants may increase their loss of carbon through root turnover, respiration or exudation. Carbon allocation to these rapid-turnover pools limits the quantity of long-term carbon storage in response to elevated CO₂ because most of the carbon is quickly returned to the atmosphere as CO₂ (refs 12, 13). Many studies have documented changes in the root biomass of plants grown under elevated CO₂ (refs 14, 15), but root turnover, exudation and respiration are difficult to quantify directly for budgets of total carbon partitioning to roots¹⁶. We have used two indirect approaches to assess carbon partitioning to roots. First, we used a mass balance model to describe the carbon stocks, partitioning and annual carbon fluxes in two annual grasslands in California after three years of exposure to elevated CO₂. Second, we used isotope labelling of reconstructed ecosystems exposed to elevated CO₂ to measure the partitioning of below-ground carbon fluxes.

Our field experimental system consists of naturally occurring

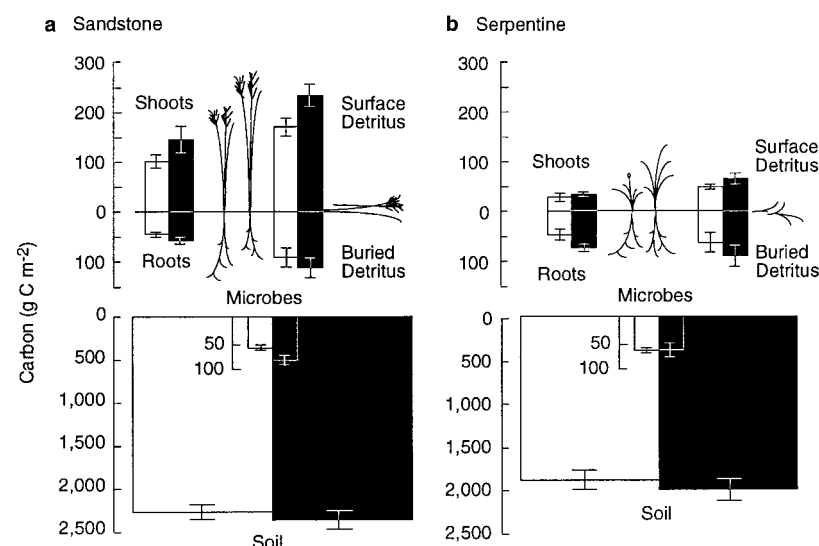


Figure 1 Carbon pools (g C m⁻², mean ± s.e.) at peak biomass in 1994 for sandstone (a) and serpentine (b) grasslands in ambient (open bars) and elevated (filled bars) CO₂.

annual grasslands in central coastal California, growing on serpentine- and sandstone-derived soils at the Jasper Ridge biological preserve of Stanford University, California (37° 24' N, 122° 14' W; elevation 150 m). The climate is Mediterranean-type, with a winter rainy season and a summer drought. The serpentine and sandstone grasslands occur adjacent to one another but differ dramatically in species composition and productivity. Introduced European annual grasses are the dominant plants on the moderately productive sandstone, whereas native forbs are dominant on the less productive serpentine grassland^{17,18}.

On each grassland, open-top chambers maintain ten ambient and ten elevated (ambient + 360 p.p.m.) CO₂-treated plots^{18,19}. We focus on results from the third growing season (1994) after establishing the CO₂ manipulation. Live shoot, live root, surface detritus, buried detritus, soil microbial and soil carbon pools were measured by destructive harvest when plants were approaching their maximum biomass during the 1994 growing season. We determined the effect of elevated CO₂ on total ecosystem carbon uptake by comparing carbon stocks in the ambient and elevated CO₂ treatments. We combined these measurements of carbon stocks with the annual carbon flux in below-ground respiration to calculate carbon partitioning to labile fractions below ground by mass balance. We also determined rates of carbon input to soil by using the ¹³C-depleted isotopic signature of the CO₂ added to the enriched plots. The added CO₂ is depleted in ¹³C ($\delta^{13}\text{C} -35\text{‰}$ versus $\delta^{13}\text{C} -8\text{‰}$ for atmospheric CO₂) because it is derived from fossil fuel. We quantified short-term carbon partitioning in below-ground respiration in a separate experiment using a ¹³CO₂ pulse label in these grassland communities grown in outdoor microcosms¹⁸. By

monitoring the rate and $\delta^{13}\text{C}$ of CO₂ released from the soil, we could distinguish CO₂ produced from root and heterotrophic sources.

In the field experiment, elevated CO₂ significantly increased carbon pools in roots, surface detritus, and soil microorganisms in the sandstone grassland, and there was a 37% increase in the total amount of carbon in these carbon pools (Fig. 1 and Table 1). Similarly, although roots were the only carbon pool (when analysed separately) to increase significantly under elevated CO₂ in the serpentine grassland (Fig. 1; one-way analysis of variance (ANOVA) $P = 0.048$), the sum of plant, detrital and microbial carbon pools increased by 25% in the serpentine grassland (Table 1). Increases in these rapidly cycling carbon pools on both grasslands (two-way ANOVA, $P = 0.002$) resulted in an increase in total ecosystem carbon (Table 1; two-way ANOVA, $P = 0.056$), even though soil carbon did not change (two-way ANOVA, $P = 0.30$)²⁰.

Significant changes in the large pool of total soil carbon (2,111 and 2,311 g C m⁻² for the serpentine and sandstone grasslands, respectively) are difficult to detect over the short time course of this experiment (discussed in detail in ref. 20). As expected, the $\delta^{13}\text{C}$ of soil carbon was lower in elevated compared to ambient CO₂ (Fig. 2), owing to the incorporation in elevated CO₂ of 266 and 303 g C m⁻² of the fossil-fuel-derived (¹³C-depleted) carbon into the serpentine and sandstone soils (0–15 cm depth), respectively, after four years²⁰. These are substantial amounts of carbon, but they represent only 13% of the soil organic carbon pool. Small increases in soil carbon can, however, lead to large increases in soil respiration if the carbon is delivered to one or more highly labile fractions in the soil²¹.

Table 1 Ecosystem carbon stocks

Ecosystem	Active C Pools* (g C m ⁻²)			Total C Pools (g C m ⁻²)		
	Ambient	Elevated	Change (%)	Ambient	Elevated	Change (%)
Sandstone	468 ± 24†	640 ± 47	37%‡	2,658 ± 85	2,868 ± 107	7.9%
Serpentine	241 ± 32	301 ± 32	25%	2,066 ± 103	2,255 ± 107	9.1%

Data are for serpentine and sandstone grasslands under ambient and elevated atmospheric CO₂ treatments.

* The sum of plant, detrital, and microbial carbon pools.

† Values are means ± standard error ($n = 9-10$).

‡ The percentage increase caused by the elevated CO₂ treatment.

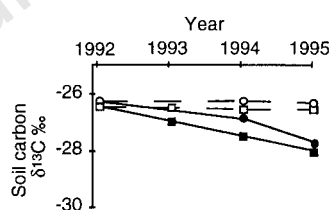


Figure 2 Soil $\delta^{13}\text{C}$ (0–15 cm) as a function of time. Soil sampling and $\delta^{13}\text{C}$ analyses are described in detail in ref. 20. Values shown are mean ± s.e. for ambient (open squares) and elevated (filled squares) CO₂ treatments on the sandstone, and ambient (open circles) and elevated (filled circles) CO₂ treatments on the serpentine grasslands. $\delta^{13}\text{C}$ declines in elevated CO₂ owing to the incorporation into soil of the ¹³C-depleted CO₂ added to these plots. Rates of carbon flux to soil were quite similar in the two grasslands, despite the two-fold greater above-ground productivity in the sandstone, underscoring the higher relative below-ground allocation in the serpentine grassland²².

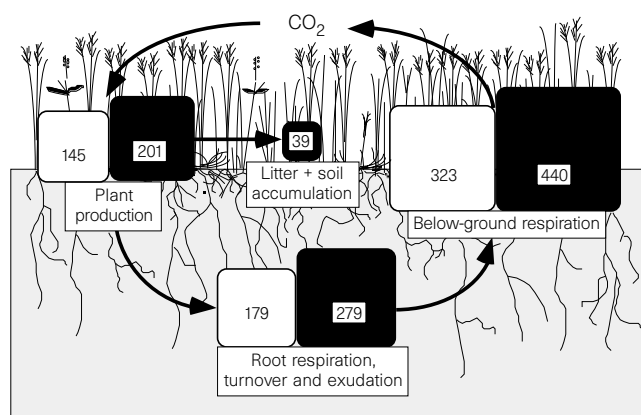


Figure 3 Annual carbon fluxes in g C m⁻² yr⁻¹ for 1994 in the sandstone grassland in ambient (white boxes) and elevated (black boxes) CO₂. Spot measurements in the field experiment and measurements over a 7-month period in the microcosm experiments show that the CO₂ stimulation of below-ground respiration in the serpentine is similar to that observed in the sandstone grassland²² (measurements in the field were too infrequent to calculate the annual flux in the serpentine).

The small stature and annual turnover of plants in these grasslands make it possible to calculate carbon allocation to labile fractions below ground by mass balance. Annual litter input to soil, excluding root exudation and turnover, can be determined as above-ground plant production plus standing root mass. Below-ground respiration includes total carbon lost through decomposition of soil organic matter and root respiration. We measured above-ground plant production, peak-season root mass, and the annual loss of carbon through below-ground respiration (measured repeatedly throughout the year in the sandstone grassland and integrated to an annual flux, as described in ref. 22). From these measurements we calculated the annual carbon input by roots, which is the sum of root respiration, intra-annual root turnover, and root exudation: $RR + T + E = BR + C_{acc} - (ANPP + RB)$, where RR , T and E are root respiration, turnover and exudation, BR is carbon loss through below-ground respiration, C_{acc} is net carbon accumulation in detritus and soil, $ANPP$ is above-ground plant production, and RB is peak-season root mass²³. By measuring RB , we can distinguish carbon allocation to standing root mass from allocation to root respiration, turnover and exudation. We assessed changes in C_{acc} caused by elevated CO_2 as the difference between the two treatments (paired by block) in litter and soil pools under ambient and elevated CO_2 , thereby avoiding the equilibrium assumption in the model of ref. 23.

Elevated CO_2 increased below-ground respiration in the sand-

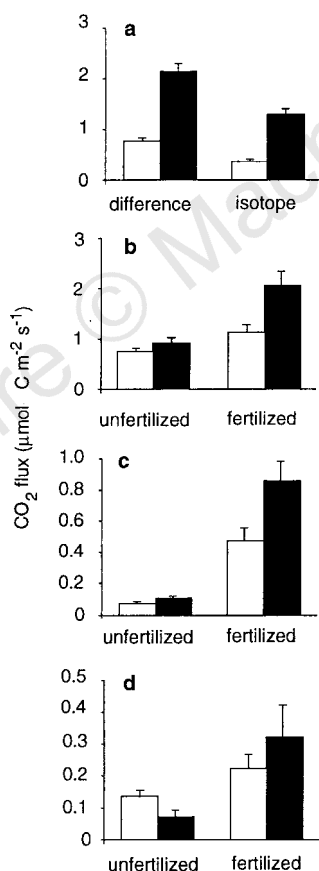


Figure 4 Partitioning below-ground respiration using the ^{13}C isotope tracer in the microcosm experiment. Values are mean $\pm$ s.e. for ambient (open bars) and elevated (filled bars) CO_2 ($n = 6-7$). Note the changes in scale for the different graphs. **a**, Root respiration rates for the unfertilized treatment only, determined either by difference or by isotope. **b**, Heterotrophic (soil microbial) respiration rates determined by laboratory incubations of root-free soil. **c**, The heterotrophic respiratory flux associated with the decomposition of rhizodeposited substrates. **d**, The heterotrophic respiratory flux associated with decomposition of above- and below-ground litter produced during the previous growing season.

stone and serpentine grasslands (Figs 3, 4)²². By the mass-balance calculation, elevated CO_2 increased root respiration, turnover and exudation by 56% in the sandstone grassland (Fig. 3), substantially more than the 25% increase in root biomass. Independent measurements using minirhizotron imaging and in-growth cores in the sandstone grassland show an increase of 0–15% in root turnover in response to elevated CO_2 (ref. 24), considerably less than the 56% increase in the sum of root respiration, turnover and exudation shown here, suggesting that root respiration and exudation increase disproportionately. The isotope experiment supported this idea, showing increased root respiration (Fig. 4a; one-way ANOVA, $P < 0.001$) and heterotrophic respiration (Fig. 4b; two-way ANOVA, $P = 0.004$), the latter resulting primarily from increased oxidation of carbon derived from roots (Fig. 4c; two-way ANOVA, $P = 0.038$). This was especially pronounced under nutrient enrichment, which increased above-ground productivity in these microcosms to levels typical of the sandstone grassland²⁵.

These findings indicate that elevated CO_2 enhances carbon partitioning to roots, a prediction that has not received wide support by studies measuring standing root biomass^{14–16}. Our results indicate that this shift in partitioning manifests primarily as increased root exudation and respiration, and is thus extremely difficult to detect, providing a possible explanation for the difficulty in accounting for all the additional carbon fixed under elevated CO_2 (refs 6, 26). All of the carbon allocated to root respiration is immediately returned to the atmosphere as CO_2 . The carbon allocated to exudation is distributed between microbial respiration and labile soil carbon²⁷. Carbon additions to labile pools in the soil can drive substantial but difficult to measure sequestration of carbon in the short term²¹. The small size and high turnover of the labile pools, however, prevents them from providing quantitatively important long-term carbon storage. Only a small portion of the labile carbon added to soil through exudation can become stabilized in soil organic matter through interactions with clays; exudates will not remain in soil owing to their chemical nature, in contrast to other more recalcitrant plant carbon constituents, such as lignin and cellulose, which may persist in soil for many years^{27,28}. Thus, compared with similar carbon additions through increased plant litter production, increased root exudation may cause relatively small increases in the carbon content of soil.

Our measurements are consistent with modelling studies that predict increased carbon uptake by grassland vegetation in response to elevated CO_2 (ref. 29), and with experimental results from tall-grass prairie where elevated CO_2 increased carbon uptake over a 34-day period³⁰. Our data provide further experimental support for increased grassland carbon uptake in response to elevated CO_2 . Yet the distribution of the extra carbon, particularly the increase in carbon allocation to labile pools below ground, suggests that the net carbon balance obtained in short-term CO_2 -enrichment experiments tend to overestimate the potential for grasslands to sequester carbon in soils in the long term.

Methods

Field experiment. Circular open-top chambers (0.65 m in diameter and 1 m tall) were established in each grassland in January 1992. A blower forces ambient air (either unsupplemented or with additional CO_2) into the lower portion of each chamber and out of the open top at a rate of $0.08 m^3 s^{-1}$, maintaining CO_2 concentrations of approximately 720 p.p.m. in the elevated treatment (for more details see refs 18, 19). Live shoot and surface litter mass were determined by harvesting a circular area (10-cm diameter) from each plot as described¹⁸ on 5 April 1994 in the serpentine and 4 May 1994 in the sandstone grassland. At the same time, 5-cm diameter by 15-cm deep cores were removed from each plot. Roots and detritus were removed from a subsample of the core by sieving (0.5 mm) followed by hand separation under a dissecting microscope; this cleaned soil was used to assess total soil and soil microbial carbon content. Roots and buried detritus (dead plant fragments, identified by colour) were recovered by washing the remainder of

the soil core. Root, detrital and soil carbon content were determined by combustion/gas chromatography (Europa Scientific); combustion/gas chromatographic determinations of shoot carbon content values from the peak biomass harvest in 1992 were used to calculate shoot carbon pools. Microbial carbon content was determined by chloroform fumigation.

Microcosm experiment. Seeds of plants representative of the serpentine and sandstone grassland communities were planted in 0.4-m diameter tubes with a 0.95-m deep column of serpentine soil in September–October 1992 and grown in open-top enclosures supplied with ambient (360 p.p.m.) or elevated (720 p.p.m.) CO₂. By the second growing season, all were similar in plant species composition and ecosystem properties, and so are combined for presentation and analysis. Nitrogen, phosphorus and potassium (20 g m⁻²) were supplied once each season as a slow-release fertilizer to half of the tubes in a factorial design. In mid-March 1993, a subset of the tubes was pulse labelled by closing the chambers and supplying 1 litre of 99% ¹³CO₂ to the chamber atmosphere during a 5-h period (8:00 to 13:00). At the height of flowering the following year (18–20 April 1994), we determined the δ¹³C of soil CO₂ (¹³CO₂ BR) collected from perforated stainless steel tubes inserted into the soil (0–15 cm depth) and the amount and δ¹³C of CO₂ produced by soil heterotrophs during 96-h laboratory incubations of root-free soil (¹³CO₂ HR). We assumed that root δ¹³C was the same as the δ¹³C of CO₂ produced by root respiration (¹³CO₂ RR) and calculated the proportional contributions from roots (y) and soil heterotrophs (1 – y) to below-ground respiration: ¹³CO₂ BR = y(¹³CO₂ RR) + (1 – y)(¹³CO₂ HR). We also calculated root respiration by difference, as below-ground respiration rates²² minus CO₂ production rates in the soil incubations. The ¹³C signal associated with the CO₂ enrichment of the atmosphere and the pulse label allowed us further to partition the heterotrophic respiration flux into three components. The soil organic carbon (SOC) contribution to heterotrophic respiration (HR) was calculated using the elevated CO₂ treatment with no pulse label, where x is the proportional contribution of SOC to HR: ¹³CO₂ HR = x(¹³C SOC) + (1 – x)(¹³C plant). Rhizodeposition and previous year's litter contributions to HR were calculated using ambient and elevated CO₂ treatments that had been pulse labelled, where y is the proportional contribution of rhizodeposition to HR and x is from above: ¹³CO₂ HR = x(¹³C SOC) + y(¹³C rhizodeposition) + (1 – x – y) (¹³C litter).

Received 10 January; accepted 9 June 1997.

- Schimel, D. et al. in *Climate Change 1995: The Science of Climate Change* (eds Houghton, J. T. et al.) 65–131 (Cambridge Univ. Press, 1996).
- Long, S. P. & Drake, B. G. in *Topics in Photosynthesis* (eds Baker, N. R. & Thomas, H.) 69–107 (Elsevier, Amsterdam, 1992).
- Broecker, W. S., Takahashi, T., Simpson, H. J. & Peng, T. H. Fate of fossil fuel carbon dioxide and the global carbon budget. *Science* **206**, 409–418 (1979).
- Gifford, R. M. Carbon dioxide and plant growth under water and light stress, implications for balancing the global carbon budget. *Search* **10**, 316–318 (1979).
- Drake, B. G. & Leadley, P. W. Canopy photosynthesis of crops and native plants exposed to long-term elevated CO₂: commissioned review. *Plant Cell Env.* **14**, 853–860 (1991).
- Canadell, J. G., Pitelka, L. F. & Ingram, J. S. I. The effects of elevated CO₂ on plant-soil carbon below ground: a synthesis. *Plant Soil* **187**, 391–400 (1996).
- Schimel, D. S. Terrestrial ecosystems and the carbon cycle. *Global Change Biol.* **1**, 77–91 (1995).
- Harrison, K., Broecker, W. & Bonani, G. A strategy for estimating the impact of CO₂ fertilisation on soil carbon storage. *Global Biogeochem. Cycles* **7**, 69–80 (1993).
- Hilbert, D. W., Larigauderie, A. & Reynolds, J. F. The influence of carbon dioxide and daily photon-flux density on optimal leaf nitrogen concentration and root:shoot ratio. *Ann. Bot.* **68**, 365–376 (1991).
- Luo, Y., Field, C. B. & Mooney, H. A. Predicting responses of photosynthesis and root fraction to elevated CO₂: Interactions among carbon, nitrogen, and growth. *Plant Cell Env.* **17**, 1195–1204 (1994).
- Field, C. B., Chapin, F. S. III, Matson, P. A. & Mooney, H. A. Responses of terrestrial ecosystems to the changing atmosphere: A resource-based approach. *Annu. Rev. Ecol. Syst.* **23**, 201–235 (1992).
- van Veen, J. A., Liljeroth, E. L., Lekkerkerk, J. A. & van de Geijn, S. C. Carbon fluxes in plant-soil systems at elevated atmospheric CO₂ levels. *Ecol. Appl.* **2**, 175–181 (1991).
- van de Geijn, S. C. & van Veen, J. A. Implications of increased carbon dioxide levels for carbon input and turnover in soils. *Vegetatio* **104/105**, 283–292 (1993).
- Stulen, I. & den Hertog, J. Root growth and functioning under atmospheric CO₂ enrichment. *Vegetatio* **104/105**, 99–115 (1993).
- Rogers, H. H., Runion, G. B. & Krupa, S. V. Plant responses to atmospheric CO₂ enrichment with emphasis on roots and the rhizosphere. *Env. Pollut.* **83**, 155–189 (1994).
- Norby, R. J. Issues and perspectives for investigating root responses to elevated atmospheric carbon dioxide. *Plant Soil* **165**, 9–20 (1994).
- Hickman, J. C. *The Jepson Manual: Higher Plants of California* (Univ. California Press, Berkeley, 1993).
- Field, C. B., Chapin, F. S. III, Chiariello, N. R., Holland, E. A. & Mooney, H. A. in *Carbon Dioxide and Terrestrial Ecosystems* (eds Koch, G. W. & Mooney, H. A.) 121–145 (Academic, San Diego, 1996).
- Jackson, R. B., Sala, O. E., Field, C. B. & Mooney, H. A. CO₂ alters water use, carbon gain, and yield for the dominant species in a natural grassland. *Oecologia* **98**, 257–262 (1994).
- Hungate, B. A., Jackson, R. B., Field, C. B. & Chapin, F. S. III Detecting changes in soil carbon in CO₂ enrichment experiments. *Plant Soil* **187**, 135–145 (1996).
- Thompson, M. V., Randerson, J. T., Malmström, C. M. & Field, C. B. Change in net primary

- production and heterotrophic respiration: how much is necessary to sustain the terrestrial sink? *Global Biogeochem. Cycles* **10**, 711–726 (1996).
- Luo, Y., Jackson, R. B., Field, C. B. & Mooney, H. A. Elevated CO₂ increases belowground respiration in California grasslands. *Oecologia* **108**, 130–137 (1996).
- Raich, J. W. & Nadelhoffer, K. J. Belowground carbon allocation in forest ecosystems: global trends. *Ecology* **70**, 1346–1354 (1989).
- Higgins, P. A. T. thesis, Stanford Univ. (1996).
- Chiariello, N. R. & Field, C. B. in *Community, Population and Evolutionary Responses to Elevated Carbon Dioxide Concentration* (eds Körner, C. & Bazzaz, F. A.) 139–175 (Academic, San Diego, 1996).
- Drake, B. G. et al. Acclimation of photosynthesis, respiration, and ecosystem carbon flux of wetland on Chesapeake Bay, Maryland, to elevated atmospheric CO₂ concentrations. *Plant Soil* **187**, 111–118 (1996).
- Parton, W. J., Schimel, D. S., Cole, C. V. & Ojima, R. Analysis of factors controlling soil organic matter levels in Great Plains grasslands. *Soil Sci. Soc. Am. J.* **51**, 1173–1179 (1987).
- Oades, J. M. The retention of organic matter in soils. *Biogeochemistry* **5**, 35–70 (1988).
- Parton, W. J. et al. Impact of climate change on grassland production and soil carbon worldwide. *Global Change Biol.* **1**, 13–22 (1995).
- Ham, J. M., Owensby, C. E., Coyne, P. I. & Bremer, D. J. Fluxes of CO₂ and water vapor from a prairie ecosystem exposed to ambient and elevated atmospheric CO₂. *Agric. For. Meteorol.* **77**, 73–93 (1995).

Acknowledgements. We thank N. Chiariello, C. Chu, G. Joel, Y. Luo, B. Mortimer, E. Nelson, J. Randerson, H. Reynolds, J. des Rosiers, S. Thayer and J. Whitbeck for contributions to the design and execution of the experiment; H. Whitted for help with experimental design and construction; P. Canadell, Z. Cardon, R. Martin and A. Townsend for assistance and advice with the ¹³CO₂ labelling, sampling and interpretation; J. Sulzmann for help with the figures; and D. Schimel for help with the isotope calculations. The Jasper Ridge CO₂ experiment is supported by grants from the US NSF to the Carnegie Institution of Washington, Stanford University and the University of California, Berkeley. B.A.H. was supported by a National Defense Science and Engineering graduate fellowship and an NSF doctoral dissertation improvement grant. R.B.J. was supported by a grant from NIGEC/DOE and a DOE distinguished postdoctoral fellowship for global change. The National Center for Atmospheric Research is sponsored by the NSF. This is publication number 1344 from the Carnegie Institution of Washington, Department of Plant Biology.

Correspondence and requests for materials should be addressed to B.A.H. (e-mail: hungate@serc.si.edu).

Net transfer of carbon between ectomycorrhizal tree species in the field

Suzanne W. Simard*, David A. Perry†, Melanie D. Jones‡, David D. Myrold§, Daniel M. Durall‡ & Randy Molina||

* Kamloops Forest Region, British Columbia Ministry of Forests, Kamloops, British Columbia V2C 2T7, Canada

† Forest Science Department and § Crop and Soil Science Department, Oregon State University, Corvallis, Oregon 97331, USA

‡ Biology Department, Okanagan University College, Kelowna, British Columbia V1V 1V7, Canada

|| United States Department of Agriculture, Forest Service, Pacific Northwest Research Station, Corvallis, Oregon 97331, USA

Different plant species can be compatible with the same species of mycorrhizal fungi^{1,2} and be connected to one another by a common mycelium^{3,4}. Transfer of carbon^{3–5}, nitrogen^{6,7} and phosphorus^{8,9} through interconnecting mycelia has been measured frequently in laboratory experiments, but it is not known whether transfer is bidirectional, whether there is a net gain by one plant over its connected partner, or whether transfer affects plant performance in the field^{10,11}. Laboratory studies using isotope tracers show that the magnitude of one-way transfer can be influenced by shading of 'receiver' plants^{3,5}, fertilization of 'donor' plants with phosphorus¹², or use of nitrogen-fixing donor plants and non-nitrogen-fixing receiver plants^{13,14}, indicating that movement may be governed by source–sink relationships. Here we use reciprocal isotope labelling in the field to demonstrate bidirectional carbon transfer between the ectomycorrhizal tree species *Betula papyrifera* and *Pseudotsuga menziesii*, resulting in net carbon gain by *P. menziesii*. *Thuja plicata* seedlings lacking ectomycorrhizae absorb small amounts of isotope, suggesting that carbon transfer between *B. papyrifera* and *P. menziesii* is primarily through the direct hyphal pathway. Net gain by *P. menziesii* seedlings represents on average 6% of carbon isotope uptake through photosynthesis. The magnitude of net transfer is influenced

by shading of *P. menziesii*, indicating that source–sink relationships regulate such carbon transfer under field conditions.

Plants within communities can be interconnected and exchange resources through a common hyphal network^{1–14}, and form guilds based on their shared mycorrhizal associates^{15–17}. Consequently, the theory that plant community dynamics operate mainly within the constraints of resource supply¹⁸ should be reformulated to consider mutualism between plants and their mycorrhizal fungi, as well as microbially mediated resource sharing^{15,17,19}.

We have studied shared compatibility of ectomycorrhizal fungi and interspecific carbon transfer in *Betula papyrifera* Marsh. and *Pseudotsuga menziesii* (Mirb.) Franco. Seven ectomycorrhizal morphotypes were common between *B. papyrifera* and *P. menziesii*, covering over 90% of their root tips, indicating a high potential for interspecific hyphal connections²⁰. We used reciprocal labelling with ¹³C and ¹⁴C in the field to examine below-ground transfer between the two tree species. To our knowledge, this is the first study in which reciprocal labelling has been used to investigate bidirectional and net carbon transfer.

We examined carbon transfer during two growing seasons, when seedlings planted in forest soil were two or three years old. Ectomycorrhizal *B. papyrifera* and *P. menziesii* and vesicular–arbuscular (VA) mycorrhizal *Thuja plicata* D. Don. were planted 0.5 m apart in three-seedling groups. These three tree species occur naturally together in the experimental area. *Thuja plicata* served as a marker for indirect isotope movement through soil pathways (that is, in respired gaseous CO₂, exudates, and sloughed root and fungal cells). Four or six weeks before labelling, *P. menziesii* seedlings in the replicate groups were exposed to three light treatments: deep shade, partial shade, and full ambient sunlight (Table 1). After the shading period, *B. papyrifera* and *P. menziesii* in each group were pulse-labelled with gaseous ¹³CO₂ or ¹⁴CO₂. This approach enabled detection of the carbon isotope that was received by one seedling from the other. Reciprocal labelling schemes were applied to replicate seedling groups in the first year to account for differences in absolute isotope fixation between ¹³C and ¹⁴C. Pulse-labelling for 2 h in full ambient light with a minimum photosynthetically active radiation level of 800 μmol m⁻² s⁻¹ was followed by a 9-day chase period under the light treatments. Seedlings were then harvested and separated into foliage, stems, coarse roots and fine roots. Ground tissue was combusted to CO₂ and analysed for ¹³C abundance by mass spectrometry and ¹⁴C by liquid scintillation.

Carbon transfer between *B. papyrifera* and *P. menziesii* occurred in both directions in the first and second year, and, although there was no net transfer in the first year, there was a net gain in carbon by *P. menziesii* in all three light intensities in the second year (Table 1 and Fig. 1). In the first year (two-year-old seedlings), isotope transfer to *P. menziesii* was balanced by reciprocal transfer to *B. papyrifera* (so there was no net transfer) in all light intensities ($P = 0.10$ for light treatment effects). Bidirectional transfer between *B. papyrifera* and *P. menziesii* represented 4% of total isotope fixed by both species together, and transfer through soil pathways to *T. plicata* was <1% of the total transfer between the ectomycorrhizal species²⁰.

In the second year (three-year-old seedlings), net transfer to *P. menziesii* and bidirectional transfer between *B. papyrifera* and *P. menziesii* represented on average 6% and 7%, respectively, of the total isotope fixed by the two tree species together, and both were over twice as great in deep shade as in partial shade or full light ($P < 0.05$; Table 1 and Fig. 1). Transfer to VA mycorrhizal *T. plicata* averaged 18% of the total transfer between *B. papyrifera* and *P. menziesii*, suggesting that most transfer between the ectomycorrhizal tree species occurred through hyphal connections²⁰. The greater effect of shading on magnitude of bidirectional transfer and the occurrence of net transfer in the second rather than the first experiment year coincided with greater root extension and potential for interplant hyphal linkages, as well as increased vigour of

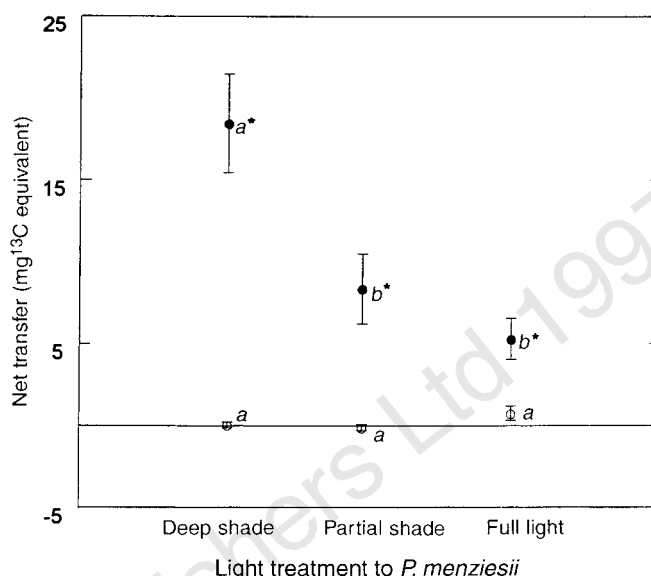


Figure 1 Mean (± s.e.) net transfer from *B. papyrifera* to *P. menziesii* in deep shade, partial shade and full ambient sunlight in the first (two-year-old seedlings, open circles) and second (three-year-old seedlings, filled circles) experiment years. Within years, means denoted by the same letter (a or b) do not differ significantly ($P < 0.05$). Asterisks indicate treatments in which net transfer was greater than zero ($P < 0.05$).

seedlings in the second year, and indicates that transfer varies according to seedling status and environmental conditions.

The greater transfer in deep shade in the second year was due to increased transfer from *B. papyrifera* to *P. menziesii*, indicating that net transfer was affected by changes in the sink strength of *P. menziesii*. Net transfer from *B. papyrifera* to *P. menziesii* was associated with whole-seedling net photosynthetic rates 1.5 and 4.3 times greater for *B. papyrifera* than for *P. menziesii* in full light and deep shade, respectively, and foliar nitrogen concentrations twice as great for *B. papyrifera* as for *P. menziesii*²⁰. These results suggest that carbon was transferred between species down carbon and nitrogen gradients. Fully developed leaves are strong sinks for nitrogen and are sources rather than sinks for carbon²¹, and amino acids rather than sugars have been shown to pass from ectomycorrhizal fungi into host plants^{22,23}. Consequently, it has been suggested that transfer may be regulated more strongly by a nitrogen gradient than a carbon gradient^{6,22}.

The 3–10% net transfer from *B. papyrifera* to *P. menziesii* measured in these experiments represents a substantial carbon gain by *P. menziesii*. It is similar to the 10% ¹⁴C transfer one way from clonal parent plants to connected ramets that has been suggested to be of sufficient magnitude to increase survival and growth of the ramets²⁴. It is also similar to the 5–15% transfer of fixed gaseous ¹⁵N₂ from *Alnus glutinosa* to *Pinus contorta* through ectomycorrhizal connections⁶, and to the 0–10% of carbon in *Cynodon dactylon* roots that was derived from *Plantago lanceolata* through VA mycorrhizal connections²⁵. Carbon transfer of this magnitude, especially if it is associated with nitrogen, would be expected to affect performance under critically harsh conditions or over the lifetime of a *P. menziesii* seedling^{10,11}.

In our experiments, the average amount of received isotope that was retranslocated from roots to foliage was 13% for *P. menziesii* and 45% for *B. papyrifera*. The amounts retranslocated from roots to foliage of receivers exceed those observed in previous ectomycorrhizal carbon-transfer experiments (usually <10%)^{5,10}, and could either supplement photosynthate or, if carbon is transferred as amino acids, the foliar nitrogen pools²³. The pattern of carbon

Table 1 Light treatment, photosynthetic rate and carbon transfer

Light treatments to <i>P. menziesii</i>	PAR incident on <i>P. menziesii</i> ($\mu\text{mol m}^{-2} \text{s}^{-1}$)*	Whole-seedling net photosynthetic rate of <i>P. menziesii</i> ($\mu\text{mol s}^{-1}$)	Bidirectional carbon transfer (mg ^{13}C equivalent)	Bidirectional transfer as % of total isotope in donor <i>B. papyrifera</i> + <i>P. menziesii</i>	Net carbon transfer as % of total isotope in donor <i>B. papyrifera</i> + <i>P. menziesii</i>
Two-year-old seedlings					
deep shade	15 $\pm$ 5 c	0.012 $\pm$ 0.002 b	0.97 $\pm$ 0.29 a	2.7	0.1
partial shade	75 $\pm$ 5 b	0.021 $\pm$ 0.005 a	1.19 $\pm$ 0.20 a	2.6	-0.3
full ambient light	175 $\pm$ 7 a	0.023 $\pm$ 0.023 a	2.13 $\pm$ 0.59 a	6.9	2.4
Three-year-old seedlings					
deep shade	60 $\pm$ 10 c	0.05 $\pm$ 0.01 b	22.06 $\pm$ 2.88 a	10.7	9.5
partial shade	250 $\pm$ 60 b	0.18 $\pm$ 0.04 a	10.87 $\pm$ 2.04 b	5.3	4.3
full ambient light	1,330 $\pm$ 50 a	0.14 $\pm$ 0.03 a	7.98 $\pm$ 1.65 b	3.9	2.7

Photosynthetically active radiation (PAR) and whole-seedling net photosynthetic rate of *P. menziesii* under three light treatments, and calculated bidirectional carbon transfer, and per cent bidirectional and net carbon transfer between *B. papyrifera* and *P. menziesii*. Values with different letters indicate significant differences among light treatments at $P < 0.05$; d.f. = 7 for two-year-old seedlings, and d.f. = 4 for three-year-old seedlings. Values are mean $\pm$ s.e.

* PAR and net photosynthetic rate were measured on a uniformly cloudy day in the first year (two-year-old seedlings) and a sunny day in the second year (three-year-old seedlings).

transfer in these ectomycorrhizal plants therefore contrasts with that reported for some VA mycorrhizal systems, where relatively little of the transferred carbon was onwardly mobilized into shoots of receiver plants, and may represent carbon restricted to fungal tissue in roots of receiver plants^{3,25,26}.

The amount of carbon exchanged between *B. papyrifera* and *P. menziesii* is indicative of a tightly linked plant–fungus–soil system. Our study extends earlier laboratory results^{3–6,10,11,24,25} to the field, providing direct evidence for both bidirectional and net carbon transfer between plant species, for the occurrence of hyphal as well as soil pathways, and for source–sink regulation of net transfer in field conditions. When *P. menziesii* seedlings are establishing and growing in the shade of illuminated *B. papyrifera* in natural, mixed communities, they may benefit directly from their association with *B. papyrifera* through supplemental gains in transferred carbon. There is a further possibility that carbon is distributed below-ground among plants across resource gradients, other than light, that affect relative photosynthetic potential within a mycorrhizally linked plant community. Such a mechanism would offer one explanation for the ability of species-rich communities to maintain productivity during drought or where nutrients are limiting²⁷. If our results reflect the magnitude of carbon transfer in natural systems, then the net competitive effect of one species on another cannot be predicted without a better understanding of interplant carbon transfer through shared mycorrhizal fungi and soil pathways. A more even distribution of carbon among plants as a result of below-ground transfer may have implications for local interspecific interactions²⁸, maintenance of biodiversity^{11,15,17}, and therefore for ecosystem productivity, stability and sustainability^{16,19,27}. □

Methods

Seedling groups. In both experiment years, each experimental unit consisted of a seedling group. In the first year (two-year-old seedlings), a 3 $\times$ 2 factorial set of treatments was replicated four times in a completely randomized design (24 experimental units). The treatments consisted of three *P. menziesii* light treatments (full ambient light, partial shade and deep shade) and two labelling schemes (^{14}C -*B. papyrifera* with ^{13}C -*P. menziesii*; ^{13}C -*B. papyrifera* with ^{14}C -*P. menziesii*). In the second year (three-year-old seedlings), each of the three light treatments was applied to five replicate groups of seedlings in a completely randomized design (15 experimental units). Owing to a shortage of replicate seedling groups, only one labelling scheme was applied in the second year: ^{14}C -*B. papyrifera* with ^{13}C -*P. menziesii*.

Light treatments. In the partial and deep shade treatments, *P. menziesii* seedlings were shaded with lightly and heavily clothed cone-shaped tents four weeks before labelling on 14–16 July 1993 (first year), and six weeks before labelling on 4–5 August 1994 (second year). The shade tents were removed from *P. menziesii* during the 2-h pulse for maximum labelling efficiency. Each *B. papyrifera* and *P. menziesii* seedling was sealed inside flexible, airtight, 5-mm thick $\times$ 60-cm wide $\times$ 90-cm tall fluoropolymer gas sampling bags (Norton Performance Plastics). The shoot of one partner seedling was then pulsed with 200 ml gaseous $^{13}\text{CO}_2$ (99%, equivalent to 107.42 mg ^{13}C), and the shoot of the

other partner was pulsed with 7.4 MBq gaseous $^{14}\text{CO}_2$ (equivalent to 52.86 μg ^{14}C , released from $\text{Na}_2^{14}\text{CO}_3$ with lactic acid).

Bidirectional and net transfer calculations. Sample $\delta^{13}\text{C}$ (‰) and ^{14}C (Bq) values were converted to excess carbon isotope (mg) for bidirectional and net transfer calculations. The conversions were based on previously described procedures^{29,30}. Conversion of $\delta^{13}\text{C}$ (‰) to the absolute isotope ratio ($^{13}\text{C}/^{12}\text{C}$) of the sample was based on the Pee Dee Belemnite (PDB) standard. Fractional abundance ($^{13}\text{C}/(^{13}\text{C} + ^{12}\text{C})$) and total carbon content (mg) of the sample were used to calculate ^{13}C content (mg) of the sample. Conversion of ^{14}C (Bq) to ^{14}C content (mg) was based on the batch-specific activity (λ) of $\text{Na}_2^{14}\text{CO}_3$ ($\lambda = 1.96 \text{ GBq mmol}^{-1}$; Amersham Canada). Excess isotope (^{13}C or ^{14}C) content (mg) of the tissue was calculated as the product of excess isotope content of the sample and tissue biomass, and excess isotope (mg) of the whole plant was determined by summing excess isotope content of the four tissue types. Bidirectional and net transfer calculations were based on whole-plant levels of excess isotope that were received from partner donor seedlings. Bidirectional transfer was the sum of isotope received by both *P. menziesii* and *B. papyrifera* in a seedling group. Conversely, net transfer was the difference between isotope received by *P. menziesii* and that received by *B. papyrifera*. Positive net transfer indicated that a greater amount of isotope was received by *P. menziesii* than by *B. papyrifera*, and negative net transfer indicated the opposite. Bidirectional and net transfer were expressed as proportions of total isotope assimilated by *P. menziesii* and *B. papyrifera* together.

Statistical analysis. Significant labelling scheme effects in the first year were detected using two factor analysis of variance (ANOVA) (α set at 0.05). For calculations of bidirectional and net carbon transfer, labelling scheme effects were first removed by applying correction factors to excess ^{14}C content (mg) on a treatment–species–tissue basis. The correction factors were the species–tissue-specific ratios of excess ^{13}C (mg) to excess ^{14}C (mg) measured in the reciprocal labelling schemes of the same light treatment. The treatment–species–tissue-specific correction factors were averaged over the four replicates per labelling scheme in the first year. The correction factors derived in the first year were applied to data collected in the second year, because only one labelling scheme was applied in the second year. Using the corrected excess ^{14}C values (analogous to excess ^{13}C equivalent values), data from the first and second years were subject to one-factor ANOVA for comparisons of net and bidirectional transfer among *P. menziesii* light treatments (7 degrees of freedom for the first year, 4 for the second). Net transfer estimates were compared with zero using *t*-tests.

Received 24 April; accepted 16 May 1997.

- Molina, R., Massicotte, H. & Trappe, J. M. in *Mycorrhizal Functioning: An Integrative Plant–Fungal Progress* (ed. Allen, M. E.) 357–423 (Chapman and Hall, New York, 1992).
- Massicotte, H. G., Molina, R., Luoma, D. L. & Smith, J. E. Biology of the ectomycorrhizal genus, *Rhizopogon*. II. Patterns of host-fungus specificity following spore inoculation of diverse hosts grown in monoculture and dual culture. *New Phytol.* **126**, 677–690 (1994).
- Francis, R. & Read, D. J. Direct transfer of carbon between plants connected by vesicular–arbuscular mycorrhizal mycelium. *Nature* **307**, 53–56 (1984).
- Brownlee, C., Duddridge, J. A., Malibari, A. & Read, D. J. The structure and function of ectomycorrhizal roots with special reference to their role in forming interplant connections and providing pathways for assimilate and water transport. *Plant Soil* **71**, 433–443 (1983).
- Finlay, R. & Read, D. J. The structure and function of the vegetative mycelium of ectomycorrhizal plants. I. Translocation of ^{14}C -labeled carbon between plants interconnected by a common mycelium. *New Phytol.* **103**, 143–156 (1986).
- Arnebrant, K., Ek, H., Finlay, R. D. & Soderstrom, B. Nitrogen translocation between *Alnus glutinosa*

- (L.) Gaertn. seedlings inoculated with *Frankia* sp. and *Pinus contorta* Dougl. ex Loud seedlings connected by a common ectomycorrhizal mycelium. *New Phytol.* **130**, 231–242 (1993).
7. Bethlenfalvay, G. J., Reyes-Solis, M. G., Camel, S. B. & Ferrera-Cerrato, R. Nutrient transfer between the root zones of soybean and maize plants connected by a common mycorrhizal mycelium. *Physiologia Plantarum* **82**, 423–432 (1991).
 8. Newman, E. I. & Eason, W. R. Rates of phosphorus transfer within and between ryegrass (*Lolium perenne*) plants. *Funct. Ecol.* **7**, 242–248 (1993).
 9. Whittingham, J. & Read, D. J. Vesicular-arbuscular mycorrhiza in natural vegetation systems III. Nutrient transfer between plants with mycorrhizal connections. *New Phytol.* **90**, 277–284 (1982).
 10. Newman, E. I. Mycorrhizal links between plants: their functioning and ecological significance. *Adv. Ecol. Res.* **18**, 243–270 (1988).
 11. Miller, S. L. & Allen, E. B. in *Mycorrhizal Functioning: An Integrative Plant–Fungal Process* (ed. Allen, M. F.) 301–332 (Chapman and Hall, New York, 1992).
 12. Ritz, K. & Newman, E. I. Nutrient transport between ryegrass plants differing in nutrient status. *Oecologia* **70**, 128–131 (1986).
 13. Frey, B. & Schüepp, H. Transfer of symbiotically fixed nitrogen from berseem (*Trifolium alexandrinum* L.) to maize via vesicular arbuscular mycorrhizal hyphae. *New Phytol.* **125**, 447–454 (1992).
 14. Ekblad, A. & Huss-Danell, K. Nitrogen fixation by *Alnus incana* and nitrogen transfer from *A. incana* to *Pinus sylvestris* influenced by macronutrients and ectomycorrhiza. *New Phytol.* **131**, 453–459 (1995).
 15. Perry, D. A., Bell, T. & Amaranthus, M. P. in *The Ecology of Mixed-Species Stands of Trees* (eds Cannell, M. G. R., Malcom, D. C. & Robertson, P. A.) 151–179 (Cambridge Univ. Press, 1992).
 16. Perry, D. A., Amaranthus, M. P., Borchers, J. G., Borchers, S. L. & Brainerd, R. E. Bootstrapping in ecosystems. *Bioscience* **39**, 230–237.
 17. Read, D. J. in *Biodiversity and Ecosystem Function* (eds Schulze, E.-D. & Mooney, H. A.) 181–209 (Springer, Berlin, 1994).
 18. Tilman, D. *Dynamics and Structure of Plant Communities* (Princeton Univ. Press, NJ, 1988).
 19. Schulze, E.-D. & Mooney, H. A. in *Biodiversity and Ecosystem Function* (eds Schulze, E.-D. & Mooney, H. A.) 497–510 (Springer, Berlin, 1994).
 20. Simard, S. W. thesis, Oregon State Univ. (1995).
 21. Pearcy, R. W. et al. Carbon gain by plants in natural environments. *Bioscience* **37**, 21–28 (1987).
 22. Smith, S. E. & Smith, F. A. Structure and function of the interfaces in biotrophic symbioses as they relate to nutrient transport. *New Phytol.* **114**, 1–38 (1990).
 23. Abuzinadah, R. A. & Read, D. J. Carbon transfer associated with assimilation of organic nitrogen sources by silver birch (*Betula pendula* Roth.). *Trees* **3**, 17–23 (1989).
 24. Alpert, P., Warembourg, F. R. & Roy, J. Transport of carbon among connected ramets of *Eichhornia crassipes* (Pontederiaceae) at normal and high levels of CO₂. *Am. J. Bot.* **78**, 1459–1466 (1991).
 25. Watkins, N. K., Fitter, A. H., Graves, J. D. & Robinson, D. Carbon transfer between C₃ and C₄ plants linked by a common mycorrhizal network, quantified using stable carbon isotopes. *Soil Biol. Biochem.* **28**, 471–477 (1996).
 26. Waters, J. R. & Borowicz, V. A. Effect of clipping, benomyl, and genet on ¹⁴C transfer between mycorrhizal plants. *Oikos* **71**, 246–252 (1994).
 27. Tilman, D., Wedin, D. & Knops, J. Productivity and sustainability influence biodiversity in grassland ecosystems. *Nature* **379**, 718–720 (1996).
 28. Perry, D. A., Margolis, H., Choquette, C., Molina, R. & Trappe, J. M. Ectomycorrhizal mediation of competition between coniferous tree species. *New Phytol.* **112**, 501–511 (1989).
 29. Boutton, T. W. in *Carbon Isotope Techniques* (eds Coleman, D. C. & Fry, B.) 155–170 (Academic, San Diego, 1991).
 30. Warembourg, F. R. & Kummerow, J. in *Carbon Isotope Techniques* (eds Coleman, D. C. & Fry, B.) 11–37 (Academic, San Diego, 1991).

Acknowledgements. We thank B. Danielson and C. Y. Li for review of the experiment designs; J. Smith and D. McKay for help with identification of ectomycorrhizal morphotypes and assistance in the greenhouse; C. Weicker and C. Gordon for assistance in the field; B. Zimonick, A. Vyse and P. G. Comeau for help and support; the staff at the Kamloops Forest Region of the British Columbia Ministry of Forests and Forestry Science Laboratory at Oregon State University for assistance; and S. Smith and R. Finlay for comments on the manuscript. This work was supported by the British Columbia Ministry of Forests and the Canada-British Columbia Partnership Agreement on Forest Resource Development (FRDA II).

Correspondence and requests for materials should be addressed to S.W.S. (e-mail: ssimard@mfor01.for.gov.bc.ca).

Dissociating prefrontal and hippocampal function in episodic memory encoding

R. J. Dolan*† & P. C. Fletcher*

* Wellcome Department of Cognitive Neurology, Institute of Neurology, Queen Square, London WC1N 3BG, UK

† Royal Free Hospital School of Medicine, Roland Hill Street, London NW3, UK

Human lesion data indicate that an intact left hippocampal formation is necessary for auditory–verbal memory¹. By contrast, functional neuroimaging has highlighted the role of the left prefrontal cortex^{2–4} but has generally failed to reveal the predicted left hippocampal activation. Here we describe an experiment involving learning category–exemplar word pairs (such as ‘dog...boxer’) in which we manipulate the novelty of either individual elements or the entire category–exemplar pairing.

We demonstrate both left medial temporal (including hippocampal) and left prefrontal activation and show that these activations are dissociable with respect to encoding demands. Left prefrontal activation is maximal with a change in category–exemplar pairings, whereas medial temporal activation is sensitive to the overall degree of novelty. Thus, left prefrontal cortex is sensitive to processes required to establish meaningful connections between a category and its exemplar, a process maximized when a previously formed connection is changed. Conversely, the left medial temporal activation reflects processes that register the overall novelty of the presented material. Our results provide striking evidence of functionally dissociable roles for the prefrontal cortex and hippocampal formation during learning of auditory–verbal material.

The role of medial temporal structures, including the hippocampal and parahippocampal formation, in episodic memory is well established, with homologous left- and right-sided structures mediating verbal and visual aspects of memory, respectively^{1,5,6}. In the monkey, lesions restricted to the hippocampus lead to striking memory impairment⁷. Memory processes can be fractionated into those relating to encoding (learning), consolidation (storage) and retrieval (recall). Functional neuroimaging, although failing to specify the role of the medial temporal cortex, has emphasized specific involvement of the left and right prefrontal cortices at encoding and retrieval respectively^{2–4}. In this experiment we sought to reconcile prefrontal and hippocampal involvement in encoding by showing that hippocampal activation reflects a response to the relative novelty of sensory data whereas prefrontal activation reflects an aspect of associative semantic processing. By novelty we refer to the relative familiarity of the verbal material in the context of the experiment; associative semantic processing refers to the formation of meaningful associations between verbal items.

The study material consisted of word-paired associates comprising a category (for example, dog) and an exemplar (for example, boxer) presented auditorily. Twelve separate perfusion scans were acquired in six male subjects while they were presented with sets of 12 paired associates. The scanning window was preceded by the presentation of two sets of 12 paired associates that created the context for the critical experimental manipulation during scanning (Fig. 1). We used a design in which category and exemplar were independent factors, with two levels related to whether an item was novel (that is, not previously presented) or old (previously presented). The study thus comprised four separate conditions representing two levels of novelty (new versus old) with respect to either category or exemplar (Fig. 1). So, for example, the initial presentations of a word pair (such as dog ... boxer) during the lead-in period might be followed, during the critical scanning period, with an entirely new pair (new–new condition: for example, food ... biscuit). Alternatively, it may be followed by a change in category (new–old condition, such as sportsman ... boxer) or a change in exemplar (old–new condition: for example, dog ... labrador). In the fourth condition, subjects heard repeat presentation of pairs presented during the lead in (old–old condition; that is, dog ... boxer). Thus, in the new–new condition, the relative emphasis is on novelty, whereas in the new–old and old–new conditions, the process emphasized is the making and breaking of semantic linkages.

Activation of the left dorsolateral prefrontal cortex (DLPFC) was sensitive to a manipulation of the association between category and exemplar or vice versa. Thus, maximal activation in this region was seen in the two conditions involving a change in category exemplar pairings. Figure 2a shows a statistical parametric map (SPM) of this activation for the contrast of the two conditions (new–old and old–new) in which there was a change in category–exemplar linkage with the condition where there was no change (old–old) ($P < 0.05$, corrected). These data are shown in Fig. 2b, where the mean adjusted activity in the left DLPFC is plotted for each of the four

experimental conditions. It can be seen that maximal activation is expressed in the two conditions involving a change in category–exemplar pairings compared to conditions in which there is no change with respect to previously established pairings. A further contrast of both conditions involving change in category exemplar pairings (new–old and old–new) with the two other conditions combined (new–new and old–old) again showed that the same left dorsolateral prefrontal cortex region is the most significant focus of activation. Additional significant activation foci were seen in the medial parietal cortex and left middle frontal gyrus. Examination of their activation profiles indicated that they represented relative decreases in the new–new condition. An activation in the left inferior parietal cortex was not predicted and does not survive correction. Table 1 lists the coordinates and associated z-scores for all of these contrasts.

By contrast, the left medial temporal cortex, including the hippocampus and parahippocampal region, showed a striking responsiveness to novelty in the presented material. Thus, maximal hippocampal activation was seen in the condition in which both category and exemplar were novel. This was significant when the condition involving a novel category and item was compared with each of the other conditions ($P < 0.001$, uncorrected). The data for the contrast of the condition where category and exemplar are new

compared to the condition when both are old is displayed as an SPM in Fig. 3a. The mean adjusted activity at the pixel of maximal significance, located in the hippocampus, for each separate condition is illustrated graphically in Fig. 3b. This shows a significant stepwise pattern of activation, with the highest values being expressed when category and exemplar are novel, intermediate when either category or exemplar are novel, and lowest when neither are novel. The associated foci of activation are shown in Table 1. To test the robustness of the left medial temporal activation further, we repeated this analysis of the hippocampal activation on each individual subject and found that in all cases, there was significant medial temporal activation ($P < 0.05$, uncorrected).

The crucial finding in this study is that there are functionally dissociable roles for the left medial temporal cortex and left prefrontal cortex during encoding of auditory–verbal material. This functional and anatomical dissociation implies that anatomically distinct brain regions have differential sensitivities to stimulus parameters and, by implication, processing requirements during encoding. These parameters may well reflect different stages in the encoding process⁸.

The importance of the human hippocampus and related medial temporal structures in auditory–verbal memory is exemplified by amnesic patients with selective lesions of this structure⁹. Activation

Lead-in list 1	Lead-in list 2	Scanning time
12 paired associates e.g. Game...Bridge Dog...Boxer	Identical to lead-in list 1	Condition 1 (New-New) New category/New exemplar e.g. Stone...Granite Cloth...Velvet Condition 2 (New-Old) New category/Old exemplar e.g. Structure...Bridge Sportsman...Boxer Condition 3 (Old-New) Old category/New exemplar e.g. Game...Football Dog...Dalmatian Condition 4 (Old-Old) Old category/Old exemplar e.g. Game...Bridge Dog...Boxer

Figure 1 The experimental design, which involved subjects hearing 12 word-paired associates while being scanned. Before scanning there was a lead-in, during which subjects had two presentations of 12 paired associates which provided the context for the experimental conditions presented during scanning. The paired associates consisted of a category and a semantically related exemplar presented at a rate of one per 4 s. There were four experimental conditions whose order was counterbalanced within and between subjects. The critical experimental conditions, each repeated three times, were defined by the degree of novelty of the category or exemplar as illustrated.

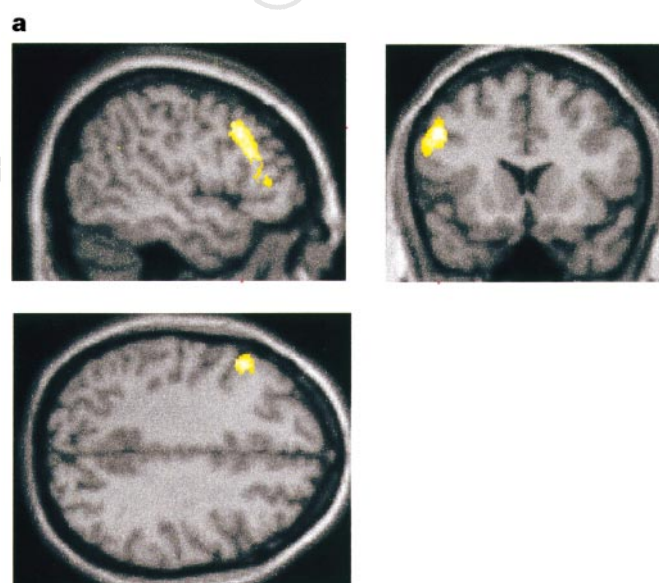
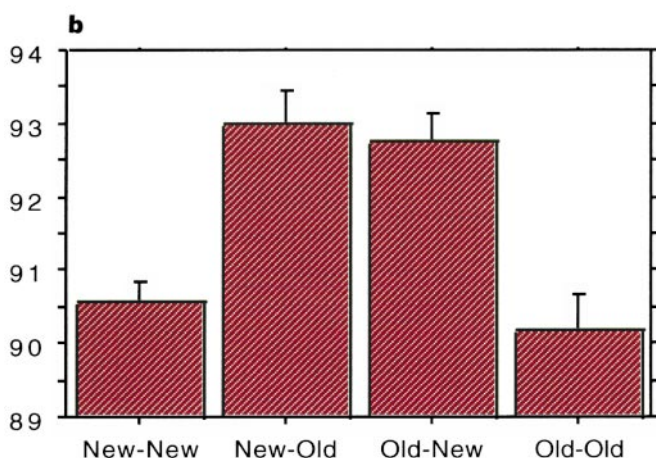


Figure 2 a, A statistical parametric map (SPM)²⁹ of a comparison between the combined 'new–old' and 'old–new' and the 'old–old' conditions, which shows greater activation in the left dorsolateral prefrontal cortex. The SPM has been rendered into standard stereotactic space and superimposed on to orthogonal sections, at voxel coordinates (x,y,z were –46, 16, 32, respectively; z score, 4.7) of a magnetic resonance image (MRI) which is itself in standard space. **b**, Regional



cerebral blood flow (rCBF) equivalents (together with standard error bars) from the most significantly activated voxel in the left DLPFC, from the comparison in **a** for each experimental condition. This voxel has identical coordinates to that in **a**. It can be seen that activation is significantly greater for the conditions involving a change in category–exemplar pairings than either of the two other conditions.

of this region has been demonstrated in an auditory-verbal paradigm that confounded encoding and retrieval¹⁰. We are aware of a single previous report of hippocampal activation in an auditory task that compared item with location encoding¹¹. Although this activation was unexpected, it suggests that processing demands at encoding other than novelty may engage the hippocampus. This is in accord with our theory that encoding involves a number of different processes that are anatomically dissociable.

In our study, when both category and exemplar are presented for the first time, the critical processing emphasized is novelty of material (with respect to the overall experimental setting) and, as predicted, robust left medial temporal activation was elicited relative to all other conditions. The involvement of medial temporal structures in novelty encoding was indicated by results from electrophysiological studies in humans and by single-cell recordings from animals^{12–14}. Neuropsychological data also indicate the absence of a characteristic electrophysiological response to novelty

in patients with hippocampal lesions¹⁵. A sensitivity to novelty is also implicit in the suggestion that hippocampal-dependent circuits maintain a template of the recent past for comparison with incoming stimuli¹⁶. Previous functional imaging studies have shown activation of the hippocampal formation during encoding of novel visual stimuli, pictures or faces^{8,17,18}. In our study, the material presented (words) was familiar, but a manipulation of the relative novelty of these words, within the experimental context engendered hippocampal activation. Note that the magnitude of hippocampal activation reflected the overall degree of novelty of the material. Our data thus support the view that the human hippocampus and related medial temporal structures are sensitive to novelty that extends beyond item novelty to include contextual novelty. This finding may explain the absence of reproducible medial temporal activation in previous functional imaging studies in that even baseline tasks that include novel material may engender hippocampal activation.

Table 1 Coordinates and z scores for activations in all reported comparisons

Comparison	Region	Coordinates	z score
New-New versus Old-Old	Left hippocampal and parahippocampal region	-16, -14, -12	3.5
		-28, -26, -20	3.8
		-42, -10, -24	3.9
New-Old + Old-New versus Old-Old	Left dorsolateral PFC	-46, 16, 32	4.6
		-46, 26, 24	4.0
New-Old + Old-New versus Old-Old + New-New	Left dorsolateral PFC	-46, 20, 30	5.3
		-46, 26, 24	4.7
	Medial parietal region (BA7)	0, -68, 50	5.8
	Left middle frontal gyrus (BA10)	-32, 58, 0	4.3
	Left inferior parietal cortex	-58, -48, 34	4.0

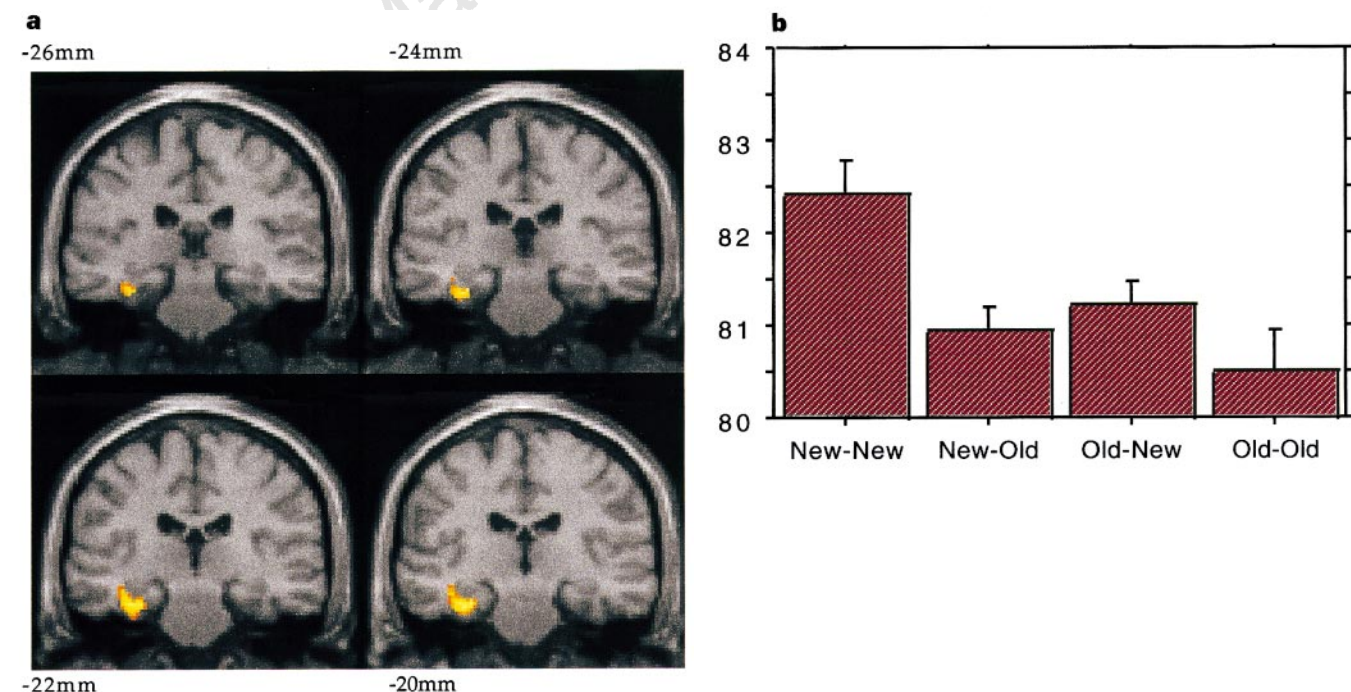


Figure 3 a, An SPM of the comparison between the 'new-new' and the 'old-old' conditions, showing significantly greater activation in the former centred on the left medial temporal cortex that incorporates the left hippocampus. The SPM has been superimposed onto 4 consecutive coronal MRI sections in the y-axis, progressing in 2-mm increments from 24 mm (top left section) to 18 mm (bottom right section) posterior to the anterior commissure. **b**, Regional cerebral blood flow equivalents (together with standard error bars) from one of the two most

significantly activated voxels (x, y, z were $-16, -14, -12$; z score, 3.5) in the left medial temporal cortex in the comparisons in **a**. The other significant voxel (x, y, z values, $-28, -26, -20$, respectively). Data from this voxel are shown for each experimental condition and indicate a stepwise pattern of activation which reflects the relative degree of novelty in the four experimental conditions: that is, the contrast of new-old and old-new activations in left medial temporal cortex were also significant ($P < 0.01$) with respect to the old-old values.

By contrast, left dorsolateral prefrontal cortex activation was maximal when either category or exemplar were novel, resulting in a new category–exemplar pairing. Activation of left prefrontal cortex is reported in a wide range of neuroimaging studies of episodic encoding^{2–4}. The foci of our activation are more dorsal than these previous reports, but the spread of activation indicates a considerable overlap. One interpretation of this activation is that it reflects retrieval of previously learned associates, but this is unlikely as cued retrieval of previously learned associates should be greatest in the contrasted old–old condition. Furthermore, retrieval of paired associates is strongly associated with right prefrontal activation³. Neuropsychological studies have indicated functional specialization in the prefrontal cortex for source, temporal sequence and strategic organizational requirements in memory^{19–22}. We found that maximal left prefrontal activation was elicited under conditions in which the critical emphasis was semantic processing necessary for the formation of new category–exemplar associations. This manipulation must elicit a degree of interference from previously encoded pairings, a phenomenon known as proactive interference. Proactive interference is a semantically mediated process and, strikingly, patients with prefrontal lesions show increased susceptibility to proactive interference^{23,24}, whereas isolated left prefrontal lesions result in an absent encoding advantage with semantic processing²⁵.

Retrieval-related prefrontal and hippocampal activations have been ascribed to effort and conscious recollection, respectively²⁶. Although we demonstrate dissociable prefrontal and hippocampal activations at encoding, we emphasize the relative nature of these differences. A hippocampal role in learning associational relations and a prefrontal sensitivity to aspects of novelty have been reported^{27,28}. Therefore, we assume that prefrontal and medial temporal activations reflect differential processing demands elicited by the experimental manipulations. In our study, prefrontal activation reflects a relative emphasis on associative semantic processing necessary in establishing and maintaining new semantic linkages in the context of already established linkages. Conversely, hippocampal activation reflects relative novelty in the study material that may engage processes that include a global integration of the sensory stimuli with the current cognitive context¹². Our data thus extend theories of hemispheric specialization for encoding by suggesting process-specific prefrontal and medial temporal activation and provide striking confirmation of the ‘novelty/encoding’ hypothesis and its predicted involvement of medial temporal structures⁸. □

Methods

Functional imaging. Six healthy male right-handed volunteers were studied using a Siemens/CPS ECAT exact HR+ (model 962) PET scanner in three-dimensional mode with a 15-cm axial field of view. Relative regional cerebral blood flow (rCBF) was measured from the distribution of radioactivity after a slow, bolus intravenous (i.v.) injection of H₂15O (8 mCi per scan, each lasting 90 s). Attenuation-corrected data were reconstructed into 63 image planes, with a resulting resolution of 6 mm at full-width-half-maximum (FWHM). For each subject, structural magnetic resonance (MR) images were obtained with a 2 T Magnetom Vision (Siemens, Germany).

Psychological tasks. Twelve rCBF measurements were made during the following four conditions: (1) novel category and novel word; (2) novel category and old word; (3) old category and novel word; (4) old category and old word. For each condition, subjects were scanned three times. While they lay with eyes closed in a quiet darkened room, stimuli were presented in an auditory verbal mode at a rate of one pair every four seconds. During the critical scanning window, subjects were presented with 12 individual word-paired associates (Fig. 1). Each of the critical experimental conditions was preceded by a run-in period in which subjects heard 2 presentations of 12 paired associates which provided the context for the presentations. Thus for example, word-paired associates might be presented twice in the run-in, whereas in the third presentation, during scanning, new exemplars might be

provided with previously presented categories as in condition (3). The order of presentation of experimental conditions was counterbalanced both within and across subjects. Subjects were instructed to attend to and memorize all the presented study material for later recall. Subjects were not informed as to when scanning occurred in relation to the presentation of the lists to ensure active attention to all the material presented. During the interscan interval, recall performance was assessed using the category as cues. For the old–new condition, subjects were asked to recall both exemplars heard at study. For the new–old condition, subjects were cued with the categories presented during the scan. These data showed recall was 95% for new–new, 83% for new–old, 73% for old–new and 93% for old–old.

Data analysis. Statistical parametric mapping (SPM96) software was used for image realignment, transformation into standard stereotactic space, smoothing and statistical analysis²⁹. All measurements per condition were averaged across subjects. State-dependent differences in global flow were co-varied out using ANCOVA. Main effects and interactions were assessed with contrasts of the adjusted task means, using *t* statistic subsequently transformed into normally distributed *z* statistic. The resulting set of *z* values constituted a statistical parametric map (SPM{*z*}), which was then thresholded at *P* < 0.001.

Received 7 March; accepted 13 May 1997.

- Scoville, W. B. & Milner, B. Loss of recent memory after bilateral hippocampal lesions. *J. Neurosurg. Psychiatr.* **20**, 11–21 (1957).
- Kapur, S. *et al.* Neuroanatomical correlates of encoding in episodic memory: levels of processing effect. *Proc. Natl Acad. Sci. USA* **91**, 2008–2011 (1994).
- Shallice, T. *et al.* Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature* **368**, 633–635 (1994).
- Fletcher, P. C. *et al.* Brain systems for encoding and retrieval of auditory-verbal memory: an *in vivo* study in humans. *Brain* **118**, 401–416 (1995).
- Smith, M. L. in *Handbook of Neuropsychology* (eds Boller, F. & Grafman, J.) 91–106 (Elsevier Science, Amsterdam, 1989).
- Dimsdale, H., Logue, V. & Piercy, M. A case of persisting impairment of recent memory following right temporal lobectomy. *Neuropsychologia* **1**, 287–298 (1964).
- Zola-Morgan, S. M. & Squire, L. R. The primate hippocampal formation: Evidence for a time-related role in memory storage. *Science* **250**, 288–290 (1990).
- Tulving, E., Markowitsch, H., Craik, F. I. M., Habib, R. & Houle, S. Novelty and familiarity activations in PET studies of memory encoding and retrieval. *Cerebr. Cort.* **6**, 71–79 (1996).
- Zola-Morgan, S., Squire, L. R. & Amaral, D. G. Human amnesia and the medial temporal region: Enduring memory impairments following a bilateral lesion limited to field CA1 of the hippocampus. *J. Neurosci.* **6**, 2950–2967 (1986).
- Grasby, P., Frith, C., Friston, K. J., Frackowiak, R. S. J. & Dolan, R. J. Activation of the human hippocampal formation during auditory-verbal long-term memory function. *Neurosci. Lett.* **163**, 185–188 (1993).
- Nyberg, J. *et al.* General and specific brain regions involved in encoding and retrieval of events: what, where, and when. *Proc. Natl Acad. Sci. USA* **93**, 11280–11285 (1996).
- Halgren, E. & Marinkovic, K. in *The Cognitive Neurosciences* (ed. Gazzinga, M. S.) 1137–1380 (MIT, Cambridge, MA, 1995).
- Fahy, F. L., Riches, I. P. & Brown, M. W. Neuronal activity related to visual recognition memory: long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Exp. Brain Res.* **93**, 457–472 (1993).
- Gershberg, F. B. & Shimamura, A. P. Impaired use of organizational strategies in free recall following frontal lobe damage. *Neuropsychologia* **13**, 1305–1333 (1995).
- Knight, R. T. Contribution of human hippocampal region to novelty detection. *Nature* **383**, 256–259 (1996).
- Martin, A., Wiggs, C., Ungerleider, L. C. & Haxby, J. C. Neural correlates of category-specific knowledge. *Nature* **379**, 649–652 (1996).
- Grady, C. L. *et al.* Age-related changes in cortical blood flow activation during visual processing of faces and location. *J. Neurosci.* **14**, 1450–1462 (1994).
- Stern, C. F. *et al.* The hippocampal formation participates in novel picture encoding: evidence from functional magnetic resonance imaging. *Proc. Natl Acad. Sci. USA* **93**, 8660–8665 (1996).
- Jetter, W., Poser, U., Freeman, R. B. & Markowitsch, H. J. A verbal long term memory deficit in frontal lobe damaged patients. *Cortex* **22**, 229–242 (1986).
- Shimamura, A. P., Janowsky, I. S. & Squire, L. R. in *Frontal Lobe Function and Dysfunction*, 173–198 (Oxford University Press, Oxford, 1991).
- Black, F. W. Cognitive deficits in patients with unilateral war related frontal lobe lesions. *J. Clin. Psychol.* **32**, 366–372 (1976).
- Schacter, D. L. Memory, amnesia and frontal-lobe dysfunction. *Psychobiology* **15**, 21–36 (1987).
- Shimamura, A. P., Jurica, P. J., Mangels, J. A., Gershberg, F. B. & Knight, R. T. Susceptibility of memory interference effects following frontal lobe damage: findings from tests of paired-associate learning. *J. Cog. Neurosci.* **7**, 144–152 (1995).
- Moscovitch, M. in *Multiple Dissociations of Function in Amnesia*, 337–370 (Hillside, NJ, 1982).
- Zatorre, R. J. & McIntee, W. Semantic encoding deficits in a case of traumatic amnesia. *Brain Cognit.* **2**, 331–345 (1983).
- Schacter, D. L., Alpert, N. M., Savage, C., Rauch, S. L. & Albert, M. S. Conscious recollection and the human hippocampal formation: evidence from positron emission tomography. *Proc. Natl Acad. Sci. USA* **93**, 321–325 (1996).
- Li, L., Miller, E. K. & Desimone, R. The representation of stimulus familiarity in the anterior inferior temporal cortex. *J. Neurophysiol.* **69**, 1918–1929 (1993).
- Raichle, M. A. *et al.* Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebr. Cort.* **4**, 8–26 (1994).
- Friston, K. J. *et al.* Statistical parametric mapping in functional imaging: a general linear approach. *Hum. Brain Map.* **2**, 189–210 (1995).

Acknowledgements. We thank our volunteers and the radiography staff at the Wellcome Department, and R. Lai for library services. R.J.D. and P.F. are supported by the Wellcome Trust.

Correspondence and requests for materials should be addressed to R.J.D.

Absence of opiate rewarding effects in mice lacking dopamine D2 receptors

Rafael Maldonado*, Adolfo Saiardi†, Olga Valverde*, Tarek A. Samad†, Bernard P. Roques* & Emiliana Borrelli†

* Département de Pharmacochimie Moléculaire et Structurale, INSERM U266, URA D1500 CNRS, Université René Descartes, 75270 Paris, France

† Institut de Génétique et de Biologie Moléculaire et Cellulaire, INSERM/CNRS/ULP, BP 163, 67404 Illkirch, C.U. de Strasbourg, France

Dopamine receptors have been implicated in the behavioural response to drugs of abuse. These responses are mediated particularly by the mesolimbic dopaminergic pathway arising in the ventral tegmental area and projecting to the limbic system. The rewarding properties of opiates¹ and the somatic expression of morphine abstinence² have been related to changes in mesolimbic dopaminergic activity that could constitute the neural substrate for opioid addiction³. These adaptive responses to repeated morphine administration have been investigated in mice with a genetic disruption of the dopaminergic D2 receptors⁴. Although the behavioural expression of morphine withdrawal was unchanged in these mice, a total suppression of morphine rewarding properties was observed in a place-preference test. This effect is specific to the drug, as mice lacking D2 receptors behaved the same as wild-type mice when food is used as reward. We conclude that the D2 receptor plays a crucial role in the motivational component of drug addiction.

Dopamine released by dopaminergic neurons in the ventral tegmental area exerts a negative control on the mesolimbic dopaminergic pathway by activating somato-dendritic D2 autoreceptors^{3,5}. The firing of these neurons is negatively regulated by GABA-mediated interneurons. This control is potentiated by stimulation of GABA receptors and is reduced by the inhibition of GABA release⁶. Opioids indirectly relieve this negative control by inhibiting the firing of these neurons. This mechanism has been proposed to mediate the rewarding effects of morphine and

heroin³. Indeed, opiate injections into this area generate self-administration⁷, potentiate rewarding brain stimulation⁸, and produce a conditioned place preference⁹. These behaviours seem to depend on an increased dopaminergic activity^{10,11}, although a dopamine-independent mechanism has also been postulated^{12–14}. Moreover, the participation of the dopaminergic mesolimbic system in physical opiate dependence has also been reported^{2,15,16}. Particularly, D2 receptors (D2Rs) within the nucleus accumbens have been shown to be involved in the expression of the somatic symptoms of opiate withdrawal².

To establish the role of dopamine D2R in such functions, we analysed the behavioural response of D2R-null mice⁴ to morphine. We have previously shown that D2R-null mice present a decrease in spontaneous locomotor activity⁴. However, D2R-deficient mice compared with their wild-type littermates ($n = 12$ per group) retain exploratory behaviour in a stressful environment (Fig. 1a). A normal habituation to the environment after three consecutive sessions is also observed (Fig. 1a). To investigate whether morphine would induce hyperlocomotion in D2R-null mice, as it does in wild-type mice¹⁷, animals were tested in locomotor-activity boxes under non-stressful conditions. Morphine administration at doses of 2 and 6 mg per kg body weight ($n = 6$ per group) induced a similar increase in horizontal locomotor activity in both wild-type and D2R-null mice (Fig. 1b).

The administration of RB 101 (100 mg per kg body weight, intraperitoneal (i.p.)), a mixed inhibitor of enkephalin-degrading enzymes¹⁸, also induced hyperactivity in both groups. Indeed, a significant increase ($P < 0.01$) in horizontal locomotor activity was observed in RB101-treated wild-type animals ($32 \pm 7\%$) and D2R-null mice ($48 \pm 15\%$) in comparison to saline-treated animals (Student's t -test: $t_{(1,12)} = 3.613$, $P < 0.01$, and $t_{(1,13)} = 2.505$, $P < 0.05$, respectively). Consequently, the motor impairment observed in D2R-null mice⁴ does not adversely affect the motor responses elicited either by exogenous or endogenous opioids. Thus the mechanisms responsible for opioid-induced enhancement of locomotion are not altered in D2R-null mice, in agreement with the previously proposed participation of D1Rs¹⁷ and of non-dopaminergic mechanisms^{19,20} in this response.

To investigate whether the reward properties elicited by morphine administration were still observable in D2R-deficient mice, wild-type and D2R-null morphine-treated mice ($n = 8$ per group) were analysed by using the place-conditioning paradigm²¹. During the

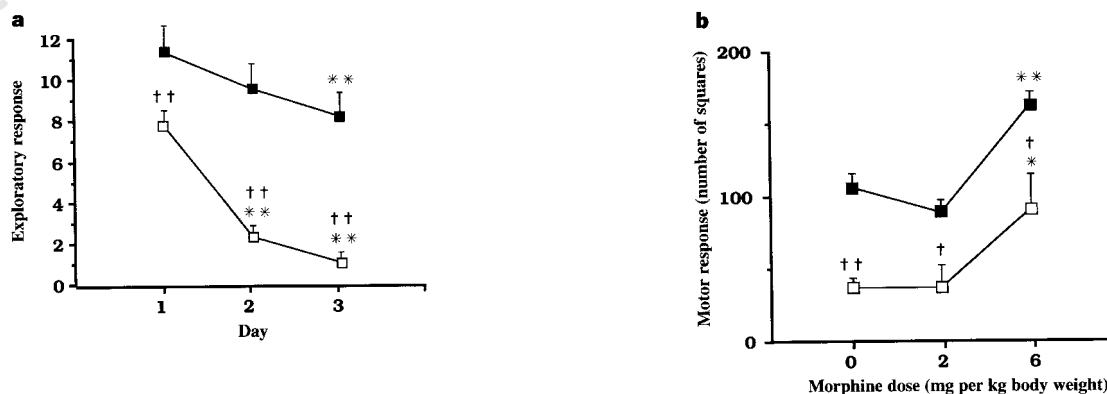


Figure 1 D2R-null mice preserve exploratory activity and locomotor response to morphine treatment. **a**, Exploratory response of D2R-null mice (open squares), and wild-type littermates (filled squares) in a stressful (highly illuminated) open field containing unknown objects. This test was performed for three consecutive days, at 24-h intervals; values represent the number of visits to the unknown objects (mean $\pm$ s.e.m.). The number of squares crossed (mean $\pm$ s.e.m.) for the wild-type mice were: day 1, 228.7 ± 10.01 ; 2, 194.8 ± 8.2 ; 3, 161 ± 11.9 ; D2R mutants, day 1, 154 ± 16.6 ; 2, 61 ± 8.4 ; 3, 47.4 ± 4 . Number of rearings for wild-

type day 1, 30 ± 1.5 ; 2, 20.8 ± 2.1 ; 3, 18.8 ± 1.8 ; for D2R mutants, day 1, 6.1 ± 1.2 ; 2, 1.9 ± 0.6 ; 3, 1.2 ± 0.4 . ** P versus same group on day 1 (Dunnett's test); †† $P < 0.01$ versus wild-type group (Student's t -test); $n = 12$ per group. **b**, Motor responses to morphine administration (mean $\pm$ s.e.m.) in D2R-null (open squares) and wild-type (filled squares) mice, in a non-stressful environment. * $P < 0.05$, ** $P < 0.01$, versus saline group with the same genotype; † $P < 0.05$, †† $P < 0.01$, versus wild-type group receiving the same treatment (Dunnett's test); $n = 6$ per group.

preconditioning phase, D2R-null and wild-type mice both visited the two compartments of the place-preference apparatus (Table 1), spending the same time in each compartment (data not shown). Two other groups of animals with the same genotypes were administered saline as control. Morphine administration (3 mg per kg body weight, subcutaneous (s.c.)) induced a clear conditioned place preference in wild-type mice, indicated by a significant increase in the time spent in the drug-associated compartment during the testing phase. This conditioned behaviour is completely absent in D2R mutant mice, which spent the same time in the morphine-designated compartment during the preconditioning and the testing phases (Fig. 2a, b). To exclude the possibility that tolerance might have arisen in D2R-null mice as a result of an endogenous higher level of enkephalins⁴, the same experiment was repeated in a new group of drug-naïve animals using a higher dose (9 mg per kg body weight s.c.) of morphine ($n = 8$ per group). In these conditions, D2R-null mice, in contrast to wild-type animals, also failed to show place preference (Fig. 2a, b). Thus the lack of place preference is not due to a developed tolerance to endogenous opioids. Motor and motivational responses to opioids have been reported to be closely related^{22,23}. However, absence of place preference is not dependent on the locomotor impairment of D2R-null mice, as these mice preserve a motor response to endogenous (RB 101 administration) and exogenous opioids (Fig. 1b); show a behaviour similar to that of wild-type mice during the preconditioning phase, including the number of visits to each compartment (Table 1); and maintain the locomotor and exploratory potential to respond to the environment (Fig. 1). Thus there is a dissociation between opioid-mediated motor and motivational responses.

The ability of D2R-null mice to show preference to other incentive stimuli was also investigated, by evaluating their condi-

tioned response to palatable food in the same place-preference paradigm. Food reward produced the same conditioned place preference in both wild-type ($n = 10$) and D2R-null mice ($n = 10$), as indicated by a significant increase in the time spent in the food-associated compartment during the testing phase (Fig. 2c, d). No place preference was observed in the control groups. Therefore the motivational deficit produced by the absence of D2Rs selectively affects morphine rewarding properties, but does not modify food-induced place preference. This suggests that there is a specific involvement of D2R in the motivational component of opioid dependence, but not on natural rewards.

It is known that μ -opioid receptors are responsible for morphine-induced place conditioned preference²⁴. We investigated whether changes in the expression of this receptor in the D2R-null background might be responsible for loss of morphine reward. Analysis of striatal and ventral mesencephalic mRNA and proteins from wild-type and D2R-null mice showed there to be a similar level of expression of this gene and of receptor binding in both genotypes (Fig. 3, and data not shown). Consequently, lack of place preference in D2R-null mice is dependent on the loss of the rewarding properties of morphine resulting from the absence of dopamine D2Rs.

We then investigated the participation of D2R in the physical dependence on opiates. The activation of D2R within the nucleus accumbens has been reported to reduce the severity of the naloxone-precipitated morphine-withdrawal syndrome, whereas its blockade in morphine-dependent animals precipitates abstinence². A severe degree of dependence was induced in wild-type and D2R-null mice by chronic administration of increasing doses of morphine. Control animals of the same genotypes were treated with saline ($n = 8$ per group). The manifestation of the somatic symptoms of withdrawal was evaluated after administration of naloxone. Chronic morphine

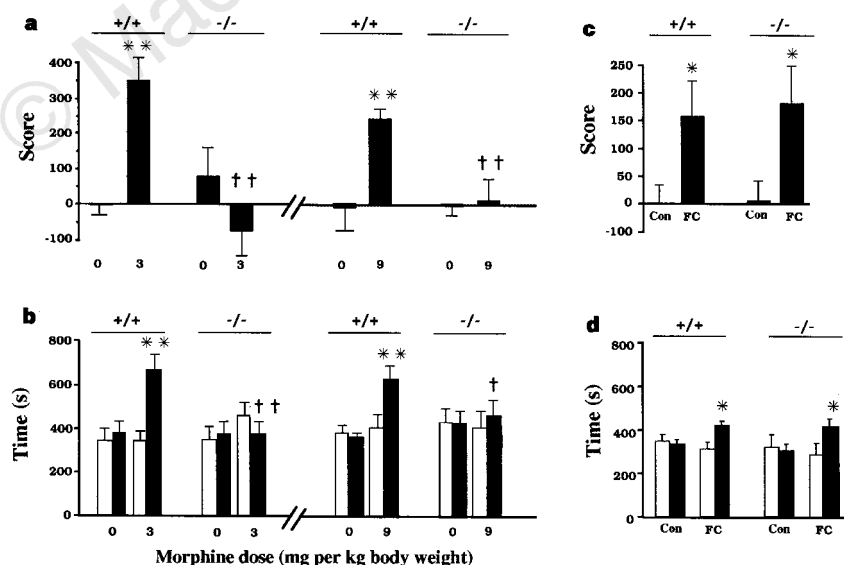


Figure 2 Lack of morphine-induced place preference in D2R-null mice. **a, b**, Conditioned place preference induced by morphine in D2R-null mice and wild-type littermates. **a**, Scores²¹ calculated as the difference between postconditioning and preconditioning time spent in the compartment associated with morphine. **b**, Time spent in the drug-associated compartment during the preconditioning (white bars) and the testing phases (black bars). (Values represent mean $\pm$ s.e.m.). ** $P < 0.01$, versus saline group with the same genotype. † $P < 0.05$, †† $P < 0.01$, versus wild-type group receiving the same treatment (Dunnett's test); $n = 8$ per group. The comparison within each group of the time spent during the preconditioning and testing phases in the morphine-associated compartment (Student's t -test) revealed a significant difference ($P < 0.01$) in wild-type mice treated with 3 and 9 mg per kg morphine (data not shown). **c, d**,

Conditioned place preference induced by food reward in D2R-null mice and their wild-type littermates. **c**, Scores (Con, control group; FC, food-conditioned group) calculated as the difference between postconditioning and preconditioning time spent in the compartment associated with food. **d**, Time spent in the food-associated compartment during the preconditioning (white bars) and testing (black bars) phases. * $P < 0.05$, versus the non-conditioned group with the same genotype (Dunnett's test); $n = 10$ per group. The comparison within each group of the time spent during the preconditioning and testing phases in the food-associated compartment (Student's t -test) also revealed a significant difference ($P < 0.05$) in food-conditioned wild-type D2R-null mice (not shown). Genotypes are indicated above the columns.

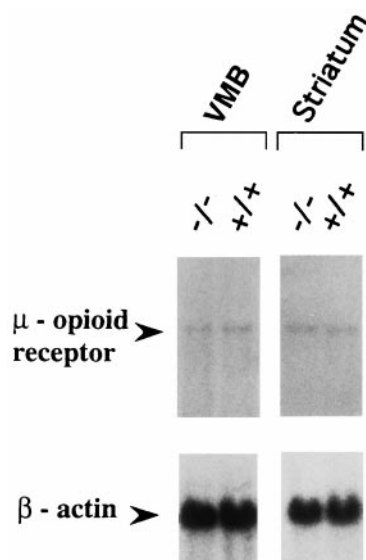


Figure 3 μ -Opioid receptor expression in D2R-null mice. Northern blot analysis of μ -opioid receptor mRNA expression in striatum and ventral midbrain (VMB) of wild type (+/+) and homozygous (-/-) D2R-null mice ($n = 3$ per group). β -Actin was used as an internal control. Bands corresponding to the μ -receptor and β -actin mRNAs are indicated.

Table 1 Frequency of visits to compartments by mice

	Wild type	Mutant
3 mg per kg morphine		
Striped compartment	26.66 $\pm$ 2.80	18.55 $\pm$ 3.95
Pointed compartment	25.77 $\pm$ 2.32	21.66 $\pm$ 3.24
Total visits	52.44 $\pm$ 4.85	40.22 $\pm$ 6.80
9 mg per kg morphine		
Striped compartment	23.36 $\pm$ 2.88	19.52 $\pm$ 4.02
Pointed compartment	32.36 $\pm$ 3.39	23.16 $\pm$ 4.76
Total visits	55.72 $\pm$ 6.27	42.68 $\pm$ 7.42

Results (mean $\pm$ s.e.m.) are for visits during the preconditioning phase.

treatment induced the classical signs of opiate exposure, such as Straub reflex and increased locomotor activity in both groups of animals (data not shown). Naloxone administration had no effect on the saline-injected groups; conversely, it strongly precipitated the behavioural symptoms of withdrawal in morphine-treated animals. No differences between D2R-null and wild-type mice were observed (Fig. 4), apart from sniffing incidence, which was significantly higher in D2R mutant mice. However, sniffing is a minor sign of abstinence and has little significance in the evaluation of withdrawal severity. Indeed, the global withdrawal scores, calculated for each animal by using a range of possible scores²⁵ from 0 to 100, were similar in mutant (64 $\pm$ 6) and wild-type mice (58 $\pm$ 5).

These data indicate that although D2 receptors in the nucleus accumbens might be involved in the modulation of physical opiate dependence², they are not required to obtain a complete manifestation of the somatic signs of opiate withdrawal. In agreement with this finding, local administration of an opiate antagonist into the nucleus accumbens did not induce any behavioural sign of withdrawal in morphine-dependent rats¹⁵. Morphine infusion into this area has also failed to produce an important degree of physical dependence¹⁶.

Previous studies of the involvement of the dopaminergic system in opiate rewarding effects have led to conflicting results^{8,12–14,26–29}. The discrepancies between these reports are probably due to the experimental conditions used, the pharmacological specificity of

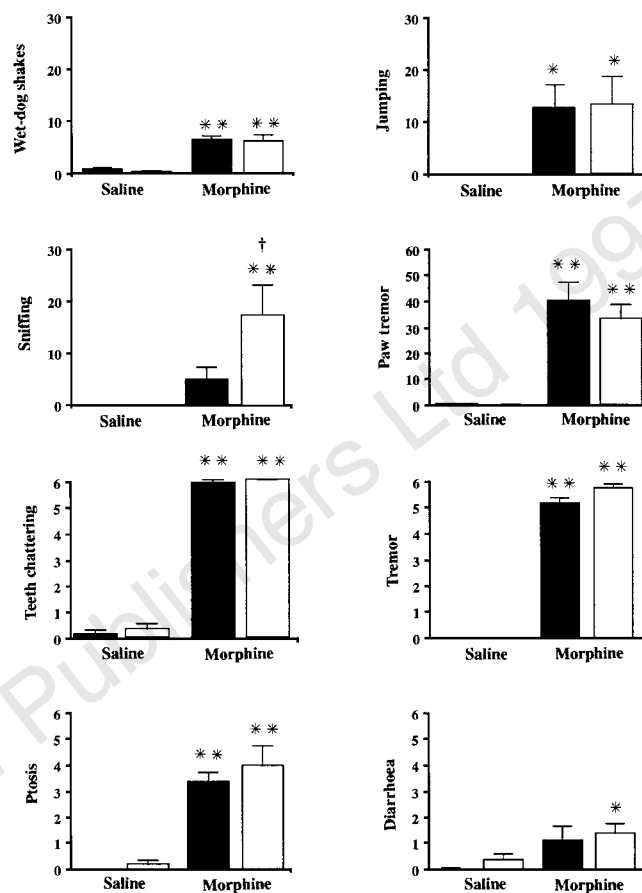


Figure 4 Analysis of morphine-induced physical dependence. Incidence of somatic signs of abstinence (mean $\pm$ s.e.m.) measured during naloxone-precipitated morphine-withdrawal syndrome in D2R-null mice (white bars) and wild-type littermates (black bars). * $P < 0.05$, ** $P < 0.01$, versus saline group with the same genotype; † $P < 0.05$, versus wild-type group receiving the same treatment (Dunnett's test); $n = 8$ per group.

the D2 antagonists, and/or the dose ranges studied. The use of D2R-null mice as a model system allows the use of intact animals in which the absence of D2R expression is obtained by genetic manipulation. Our results have unambiguously demonstrated a dissociation between the mechanisms involved in the rewarding properties of opiates, which are related to their addictive capacities, and the somatic signs of the naloxone-precipitated withdrawal syndrome, which reveal the physical component of opiate dependence. Thus dopaminergic D2 receptors are crucial for the rewarding effects of opiates, but not for the development of physical opiate dependence or the expression of some acute responses to morphine, such as hypermotility. We have also demonstrated a selective physiological involvement of D2Rs in some forms of drug-mediated reward that is different from that produced using other rewarding stimuli. □

Methods

The D2R (-/-) mice and wild-type (+/+) littermates were from the mating of heterozygous D2R-null mice⁴. Morphine-HCl and naloxone-HCl were from Sigma, and RB101 was produced as described¹⁸. All behavioural and pharmacological experiments were performed in a soundproof room by an observer who did not know either the genotype of the mice or the pharmacological treatment.

Open field test. For stressful conditions, a large white rectangular box was used (70-cm wide, 90-cm long and 60-cm high) that was brightly illuminated from the top (500 lux) to present four unknown objects, placed 20 cm from

each corner. A total of 63 squares (10×10 cm) were drawn with black lines on the white floor of the field. The presence of six events was quantified during an observation period of 5 min: squares crossed, rearings, grooming bouts, defecation boli, urination events and visits to the unknown objects. Values were analysed by two-way analysis of variance (ANOVA) with repeated measures (time, within subjects; genotype, between subjects). After main effect of ANOVA, individual comparisons within subjects were made by the two-tailed Dunnett's test, and between subjects by the two-tailed Student's *t*-test.

Locomotor activity boxes. In non-stressful conditions, locomotion was evaluated in locomotor activity boxes 10 min after saline or morphine-HCl (2 and 6 mg per kg body weight, s.c.) administration in a transparent rectangular box ($30 \times 26 \times 30$ cm) with a metallic grid floor and moderate illumination (50 lux). Mice were not habituated to the test box. Animal displacements were measured by drawing 12 squares on the floor. The number of squares crossed and the number of rearings were counted during a period of 10 min. Values were analysed by two-way ANOVA (genotype and treatment) between subjects. After main effect of ANOVA, individual comparisons between groups were made by two-tailed Dunnett's test.

Place-preference paradigm. Tests were performed as previously reported²¹, except for the conditioning time (18 min) and the size of compartments ($15 \times 15 \times 15$ cm). One compartment of the place-preference apparatus is randomly chosen to be paired to drug administration and the other to saline. The drug-administered compartment could be either the most or the least preferred. Care was taken to ensure that treatments were counterbalanced as closely as possible between compartments. Experiments were conducted during the light phase of the dark/light cycle. The place-preference conditioning schedule consisted of three phases. (1) Preconditioning phase, during which mice were placed in the middle of the neutral area and their location was recorded for the 18 min. After the session, animals were randomized to be paired to drug or vehicle administration and to be assigned to a compartment. (2) Conditioning phase, during which mice were treated for 6 consecutive days with alternate injections of morphine-HCl (3 or 9 mg per kg body weight, s.c.) or saline. Doors matching the walls of the compartment allowed the confinement of the mice for 20 min immediately after morphine or saline injections. Mice received the drug on days 1, 3 and 5, and vehicle on days 2, 4 and 6. Control animals received vehicle every day. (3) Testing phase, which was conducted exactly as the preconditioning phase (free access to each compartment for 18 min, 24 h after the final conditioning session). Food reward was tested in animals with limited access to food (normal mouse food plus sucrose) for 5 days before the test. The same procedure was used as for morphine, with the exception that mice conditioned to food had free access to it (normal mouse food plus sucrose) in the confined compartment on days 1, 3 and 5, and had no access to food in the other compartment on days 2, 4 and 6. Control animals had no access to food in the assigned compartments during this phase (days 1–6). A fixed amount of food equivalent to 12% of mouse body weight was applied per day, as previously reported³⁰. Animals were maintained on the same diet throughout the experimental period. Data were analysed by two-way ANOVA (genotype and treatment) between subjects. Individual comparisons were made in each case by the two-tailed Dunnett's test, after main effect of ANOVA.

Analysis of abstinence somatic signs. Opiate dependence was induced in mice by repeated i.p. injection of morphine-HCl, at an interval of 12 h, for 6 days. The morphine-HCl dose was progressively increased as follows: day 1, 10 mg per kg body weight; 2, 20 mg per kg; 3, 30 mg per kg; 4, 40 mg per kg; 5, 50 mg per kg; 6 (only one injection in the morning), 50 mg per kg. Control mice were treated with saline under the same conditions. Withdrawal was precipitated only once in each animal by injecting naloxone-HCl (1 mg per kg, s.c.) 2 h after the last morphine-HCl administration. Mice were placed individually into test chambers consisting of transparent round plastic boxes (30 cm diameter, 50 cm height) with a white floor 30 min before naloxone-HCl injection. During the 15 min before naloxone-HCl injection, mice were observed to verify the presence of normal behaviour. Somatic signs of withdrawal were evaluated immediately after the injection of the opiate antagonist during a period of 30 min. The number of wet-dog shakes, jumping, sniffing and paw tremor were counted. Teeth chattering, diarrhoea, tremor and ptosis were evaluated over periods of 5 min, one point being given for the

presence of each sign during each period. The number of periods showing the sign was then counted (maximum score of 6). Values were analysed by two-way ANOVA (genotype and treatment) between subjects. Individual comparisons were made by the two-tailed Dunnett's test, after main effect of ANOVA.

Northern blot analysis. Total RNA from striatum and ventral midbrain was purified by the lithium chloride method, and 10 µg RNA was analysed by northern blot. The ³²P random primed probe used to reveal the µ-opioid receptor mRNA corresponds to the 870 base-pair fragment spanning base pairs 785 to 1653 of the mouse µ-opioid receptor cDNA²⁴.

Received 4 November 1996; accepted 9 May 1997.

- Wise, R. A. & Bozarth, M. A. A psychomotor stimulant theory of addiction. *Psychol. Rev.* **97**, 469–492 (1987).
- Harris, G. C. & Aston-Jones, G. Involvement of D2 dopamine receptors in the nucleus accumbens in the opiate withdrawal syndrome. *Nature* **371**, 155–157 (1994).
- Koob, G. F. Drugs of abuse: anatomy, pharmacology and function of reward pathways. *Trends Pharmacol. Sci.* **13**, 177–184 (1992).
- Baik, J. H. et al. Parkinsonian-like locomotor impairment in mice lacking dopamine D2 receptors. *Nature* **377**, 424–428 (1995).
- Di Chiara, G. & Imperato, A. Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. *Proc. Natl Acad. Sci. USA* **85**, 5274–5278 (1988).
- Wise, R. A. Neurobiology of addiction. *Curr. Opin. Neurobiol.* **6**, 243–251 (1996).
- Bozarth, M. A. & Wise, R. A. Intracranial self-administration of morphine into the ventral tegmental area in rats. *Life Sci.* **28**, 551–555 (1981).
- Broekkamp, C. L., Phillips, A. S. & Cools, A. R. Facilitation of self-stimulation behavior following intracerebral microinjections of opioids into the ventral tegmental area. *Pharmacol. Biochem. Behav.* **11**, 289–295 (1979).
- Phillips, A. C. & Le Piane, E. G. Reinforcing effects of morphine microinjection into the ventral tegmental area. *Pharmacol. Biochem. Behav.* **12**, 965–968 (1980).
- Joyce, E. M. & Iversen, S. D. The effect of morphine applied locally to mesencephalic dopamine cell bodies on spontaneous motor activity in the rat. *Neurosci. Lett.* **14**, 207–212 (1979).
- Kalivas, P. W. & Richardson-Carlson, R. Endogenous enkephalin modulation of dopamine neurons in ventral tegmental area. *Am. J. Physiol.* **251**, R243–R249 (1986).
- Ettenberg, A., Pettit, H. O., Bloom, F. E. & Koob, G. F. Heroin and cocaine intravenous self-administration in rats: mediation by separate neural systems. *Psychopharmacology* **78**, 204–209 (1982).
- Pettit, H. O., Ettenberg, A., Bloom, F. E. & Koob, G. F. Destruction of dopamine in the nucleus accumbens selectively attenuates cocaine but not heroin self-administration in rats. *Psychopharmacology* **84**, 167–173 (1984).
- Borg, P. J. & Taylor, D. A. Voluntary oral morphine self-administration in rats: effect of haloperidol or ondansetron. *Pharmacol. Biochem. Behav.* **47**, 633–646 (1994).
- Maldonado, R., Stinus, L., Gold, L. & Koob, G. F. Role of different brain structures in the expression of the physical morphine withdrawal syndrome. *J. Pharmacol. Exp. Ther.* **261**, 669–677 (1992).
- Bozarth, M. A. Physical dependence produced by central morphine infusions: an anatomical mapping study. *Neurosci. Biobehav. Rev.* **18**, 373–383 (1994).
- Longoni, R., Spina, L. & Di Chiara, G. Dopaminergic D-1 receptors: essential role in morphine-induced hypermotility. *Psychopharmacology* **93**, 401–402 (1987).
- Fournié-Zaluski, M. C. et al. "Mixed inhibitor-prodrug" as a new approach toward systemically active inhibitors of enkephalin-degrading enzymes. *J. Med. Chem.* **35**, 2473–2481 (1992).
- Kalivas, P. W., Widerlow, E., Stanley, D., Breese, G. & Prange, A. J. Enkephalin action on the mesolimbic system: a dopamine-dependent and a dopamine-independent increase in locomotor activity. *J. Pharmacol. Exp. Ther.* **227**, 229–237 (1983).
- Daugé, V., Rossignol, P. & Roques, B. P. Blockade of dopamine receptors reverses the behavioral effects of endogenous enkephalins in the nucleus caudatus but not in the nucleus accumbens: differential involvement of delta and mu opioid receptors. *Psychopharmacology* **99**, 168–175 (1989).
- Valverde, O., Fournié-Zaluski, M. C., Roques, B. P. & Maldonado, R. The CCK₈ antagonist PD-134,308 facilitates rewarding effects of endogenous enkephalins but does not induce place preference in rats. *Psychopharmacology* **123**, 119–126 (1995).
- Salamone, J. D. The involvement of nucleus accumbens dopamine in appetitive and aversive motivation. *Behav. Brain Res.* **61**, 117–133 (1994).
- Salamone, J. D. The behavioral neurochemistry of motivation: methodological and conceptual issues in studies of the dynamic activity of nucleus accumbens dopamine. *J. Neurosci. Methods* **64**, 137–149 (1996).
- Matthes, H. W. D. et al. Loss of morphine-induced analgesia, reward effect and withdrawal symptoms in mice lacking the µ-opioid-receptor gene. *Nature* **383**, 819–823 (1996).
- Maldonado, R., Negus, S. & Koob, G. F. Precipitation of morphine withdrawal syndrome in rats by administration of µ-, delta- and kappa-selective opioid antagonists. *Neuropharmacology* **31**, 1231–1241 (1992).
- Spyraki, C., Fibiger, H. C. & Phillips, A. G. Attenuation of heroin reward in rats by disruption of the mesolimbic dopamine system. *Psychopharmacology* **79**, 278–283 (1983).
- Mackey, W. B. & van der Kooy, D. Neuroleptics block the positive reinforcing effects of amphetamine but not of morphine as measured by place conditioning. *Pharmacol. Biochem. Behav.* **22**, 101–105 (1985).
- Stinus, L. et al. Chronic flupentixol treatment potentiates the reinforcing properties of systemic heroin administration. *Biol. Psychiat.* **26**, 363–371 (1989).
- Smith, J. E., Guerin, G. F., Co, C., Barr, T. S. & Lane, J. D. Effects of 6-OHDA lesions of the central medial nucleus accumbens on rat intravenous morphine self-administration. *Pharmacol. Biochem. Behav.* **23**, 843–849 (1985).
- Elmer, G. I., Pieper, J. O., Goldberg, S. R. & George F. R. Opioid operant self-administration, analgesia, stimulation and respiratory depression in µ-deficient mice. *Psychopharmacology* **117**, 23–31 (1995).

Acknowledgements. We acknowledge B. Kieffer for the gift of µ-opioid cDNA. A.S. and O.V. were recipient of fellowships from the European Community and T.A.S. from the Centre Nationale de la Recherche Scientifique (CNRS). This work was supported by funds from the Ministère de la Recherche to R.M., from the Association pour la Recherche sur le Cancer to R.M. and to E.B., and from the Institut National de la Santé et de la Recherche Médicale, CNRS et Centre Hospitalier Universitaire Régional to E.B.

Correspondence and requests for materials should be addressed to E.B. (e-mail: eb@titus.u-stras.fr).

Rab3A is essential for mossy fibre long-term potentiation in the hippocampus

Pablo E. Castillo^{*†}, Roger Janz[‡], Thomas C. Südhof[‡], Thanos Tzounopoulos[§], Robert C. Malenka^{§||} & Roger A. Nicoll^{*||}

Departments of ^{*}Cellular and Molecular Pharmacology, [§]Psychiatry and ^{||}Physiology, University of California, San Francisco, California 94143, USA
[‡]Department of Molecular Genetics and Howard Hughes Medical Institute, The University of Texas Southwestern Medical Center, Dallas, Texas 75235, USA
[†]Present address: Departamento de Fisiología, Facultad de Medicina, Gral Flores 2125, Montevideo, Uruguay.

Repetitive activation of excitatory synapses in the central nervous system results in a long-lasting increase in synaptic transmission called long-term potentiation (LTP). It is generally believed that this synaptic plasticity may underlie certain forms of learning and memory. LTP at most synapses involves the activation of the NMDA (*N*-methyl-D-aspartate) subtype of glutamate receptor, but LTP at hippocampal mossy fibre synapses is independent of NMDA receptors and has a component that is induced and expressed presynaptically¹. It appears to be triggered by a rise in presynaptic Ca²⁺ (refs 2, 3), and requires the activation of protein kinase A^{4–6}, which leads to an increased release of glutamate^{3,7–10}. A great deal is known about the biochemical steps involved in the vesicular release of transmitter^{11–13}, but none of these steps has been directly implicated in long-term synaptic plasticity. Here we show that, although a variety of short-term plasticities are normal, LTP at mossy fibre synapses is abolished in mice lacking the synaptic vesicle protein Rab3A.

In searching for membrane fusion proteins that might be involved in mossy fibre LTP, a starting point is to look at those proteins that are substrates for protein kinase A (PKA). The most thoroughly studied synaptic proteins that are phosphorylated by

PKA are the synapsins¹⁴. However, previous studies on mutant mice in which both synapsin I and II were deleted showed normal LTP¹⁵. Another vesicle-associated protein, rabphilin-3, is also a substrate for PKA¹⁶. Rabphilin exerts its effects by binding to Rab3 proteins (specifically Rab3A and Rab3C at synapses), which are small, GTP-binding proteins that are thought to attach reversibly to the membrane of the synaptic vesicle¹³. Furthermore, a second putative Rab3-effector protein, called Rim, was recently discovered¹⁷ that also contains PKA phosphorylation consensus sequences. This suggests that Rab3A and Rab3C may have multiple targets at the synapse that are potentially regulated by PKA. Previous electrophysiological studies in the CA1 region of the hippocampus from mice lacking the Rab3A protein showed that, although most synaptic parameters were normal, there is increased synaptic depression after short trains of repetitive stimulation^{1,18}. Here we consider the physiological properties of mossy fibre synapses that exhibit short- and long-term plasticities that are profoundly different from those previously studied in the brain.

Initially we used highly sensitive immunoperoxidase labelling¹⁹ to examine the anatomy of the mossy fibre pathway in Rab3A-knockout mice and in littermates. Monoclonal antibodies specific for Rab3A (C142.2) or cross reactive with all Rab3 isoforms in brain (Rab3A, 3B and 3C; C142.1) and polyclonal antibodies to Rab3C were used to evaluate Rab3 proteins in the hippocampus (Fig. 1 and data not shown). Immunoperoxidase staining failed to reveal the presence of Rab3C or Rab3B in the stratum lucidum, the termination zone of the mossy fibres. In contrast, high levels of Rab3C were observed in the rest of the hippocampus, especially in the dentate gyrus, which is the termination zone of fibres from the entorhinal cortex. Thus Rab3A is the only detectable Rab3 isoform in the mossy fibre terminals, making this synapse an excellent system in which to study the effect of the Rab3A knockout in the absence of possibly redundant Rab3 isoforms. No induction of other Rab3 proteins or changes in the structure of the mossy fibre layer was observed in the Rab3A-knockout mice. When synaptophysin/synaptophysin II was used as a marker for mossy fibre terminals²⁰, no changes in the intensity of staining or the structure of stratum lucidum was apparent (Fig. 1). Together with the previously characterized lack of ultrastructural changes in Rab3A-knockout mice¹⁸, these data

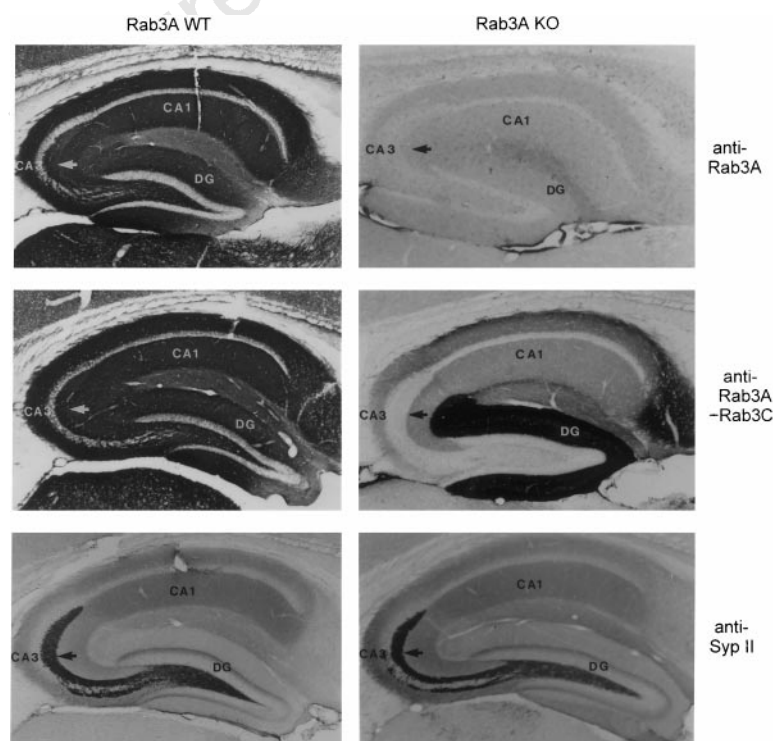


Figure 1 Immunohistochemical analysis of the Rab3 proteins in hippocampus. Hippocampal sagittal sections of wild-type (WT) and Rab3A-deficient (knockout, KO) mice were immunostained using antibodies specific for Rab3A, Rab3A/B/C, and synaptophysin II/synaptophysin. The bound antibodies were visualized using the peroxidase method with diaminobenzidine and heavy-metal enhancement¹⁹. The dentate gyrus (DG) and the CA1 and CA3 regions are labelled. The arrow marks the mossy fibre bundle. Using the highly sensitive peroxidase method, no immunoreactivity for Rab3A, B or C was detected in Rab3A-deficient mice, whereas other parts show low levels (CA1) or high levels (dentate gyrus) of Rab3B/C expression. The absence of Rab3C immunoreactivity in the mossy fibre pathway was confirmed using a Rab3C-specific polyclonal antibody (data not shown). The labelling of synaptophysin II, which is strongly expressed in the mossy fibre bundle²⁰, demonstrates that this structure is normally developed in the Rab3A-deficient mice.

suggest that the mossy fibre pathway in these mice is structurally normal but lacks other Rab3 proteins, making it uniquely suited for the study of the effects of the Rab3A knockout on synaptic function and plasticity.

Although no anatomical abnormalities could be detected, it is possible that the Rab3A deletion prevents the release of transmitter from mossy fibre synapses. Testing the functional integrity of the mossy fibre pathway is not trivial as, in the absence of functional mossy fibres, activation of neighbouring associational/commissural fibres would still be expected to elicit a response. Thus a method was required to distinguish mossy fibre synaptic responses from non-mossy fibre synaptic responses. We have found, in agreement with previous observations^{21,22}, that the mossy fibre synapses in the mouse possess inhibitory presynaptic metabotropic glutamate receptors (mGluRs) of the group 2 subtype, whereas associational/commissural fibres do not have these receptors. Thus application of the group 2-selective agonist L-CCG1 (10 μ M) had no effect on the responses to associational/commissural stimulation but completely blocked the response evoked by a stimulating electrode placed in the dentate granule cell layer (Fig. 2a). The magnitude of the inhibition induced by L-CCG1 was the same in wild-type and Rab3A-deficient mice (Fig. 2b), indicating that the mossy fibres are present and functional in the knockout animal.

We next looked for possible effects on two forms of short-term synaptic plasticity of knocking out Rab3A. When a synapse is activated twice at short intervals, the size of the response to the second stimulus is increased, as there is an increase in transmitter release²³. This is referred to as paired-pulse facilitation (PPF). In the mossy fibres, tested at a 40-ms interstimulus interval with field-potential recording, PPF exhibited no significant change in the knockout (wild type, 2.7 ± 0.3 , $n = 15$; knockout, 3.0 ± 0.4 , $n = 14$). Because PPF in the CA1 region is enhanced in the Rab3A-deficient mouse²⁴, especially at short intervals, we repeated our studies on mossy fibre synapses using the whole-cell voltage-clamp technique. No difference in PPF was found either at 20 or

40 ms (Fig. 3a). As well as exhibiting robust PPF, mossy fibre responses are also extremely sensitive to the frequency of stimulation, such that modest increases in stimulation frequency cause a large increase in transmitter release^{25,26}. This phenomenon, referred to as frequency facilitation²³, was unaffected in the knockout mouse (Fig. 3b). Thus two different forms of short-term synaptic plasticity were unaffected by the deletion of Rab3A. Both of these forms of plasticity are very sensitive to the probability of transmitter release, consistent with the idea that the probability of release is not markedly altered in the knockout mice. However, our experiments do not address the possibility that the amount of transmitter release, once release occurs, is increased or decreased in the knockout mice.

We next examined the increase in synaptic transmission evoked by raising the levels of cyclic AMP by applying forskolin. This enhancement of mossy fibre transmission is primarily the result of an increased release of neurotransmitter^{4,5}, and was also unaffected in the knockout (Fig. 3c). To confirm that the forskolin-induced enhancement was due to an elevation in cAMP, we carried out two further experiments. First, we examined the effects of Sp-8-Br-CPT-cAMPS, a membrane-permeant analogue of cAMP, on mossy fibre responses. The increase in transmission caused by Sp-8-Br-CPT-cAMPS was similar in the wild-type ($179 \pm 15\%$ of control, $n = 3$) and the knockout mice ($187 \pm 12\%$ of control, $n = 4$). Second, we found that an analogue of forskolin, 1,9-dideoxyforskolin, which does not activate adenylyl cyclase but shares some of the other effects of forskolin²⁷, had no effect on mossy fibre responses ($n = 3$).

Finally, we tested for possible effects of the knockout on mossy fibre LTP. Tetanic stimulation (given in the presence of 100 μ M D-AP5) elicited robust mossy fibre LTP in wild-type mice. In marked contrast, LTP was essentially absent in the Rab3A-deficient mice (Fig. 4). The potentiation immediately after the tetanus, however, did not seem to be altered in the knockout. In addition, the size of the synaptic response at the end of the tetanus was similar in the wild-type ($46 \pm 20\%$ of pre-tetanus response, $n = 6$) and knockout mice ($36 \pm 18\%$, $n = 6$) (not shown). These results suggest that, on

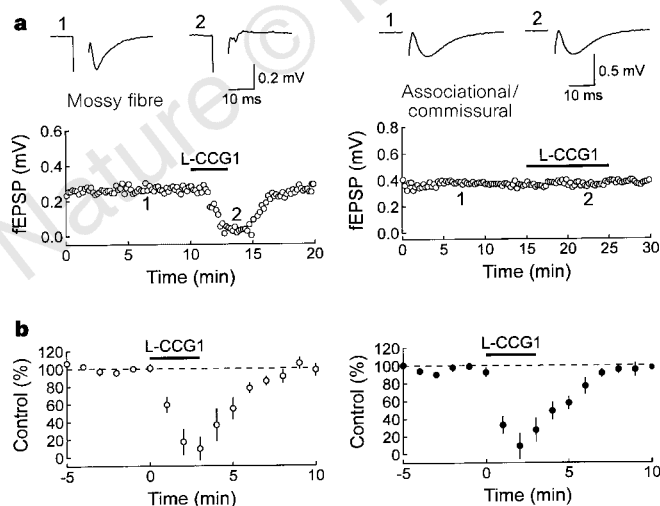


Figure 2 Mossy fibre synaptic inputs are normal in Rab3A-deficient mice. **a**, L-CCG1 (10 μ M) bath application for 3 min abolishes mossy fibre synaptic responses (left); longer application (10 min) has no effect on associational-commissural responses (right). **b**, Inhibitory effects of L-CCG1 were identical in wild-type (left, open circles; 4 animals, 6 slices) and Rab3A-mutant mice (right, filled circles; 4 animals, 7 slices), indicating that functional mossy fibre inputs are intact in these mice. fEPSP, field excitatory postsynaptic potential.

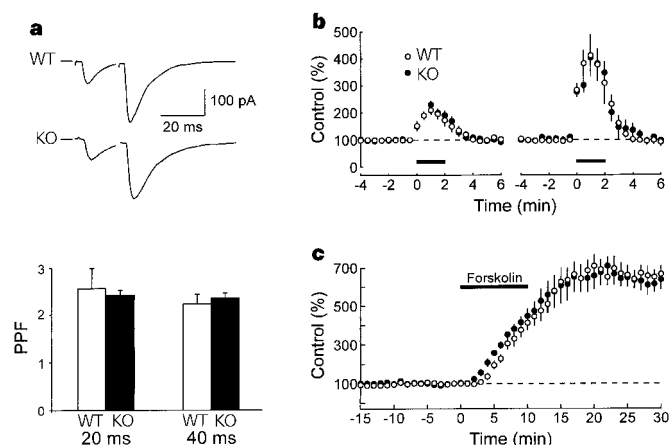


Figure 3 Synaptic transmission in Rab3A-deficient mice. **a**, Paired-pulse facilitation (PPF) measured at intervals of 20 and 40 ms was compared in normal and mutant mice using whole-cell voltage-clamp techniques. Top, typical responses from wild-type (WT) and Rab3A-knockout (KO) mice. Bottom, magnitude of PPF (mean $\pm$ s.e.m.) at 20 and 40 ms is summarized for wild-type (white bars; 3 animals, 8 cells) and knockout mice (black bars; 3 animals, 9 cells). **b**, Frequency facilitation was explored with field-potential recording in wild-type (open circles; 5 animals, 8 slices) and mutant mice (filled circles; 6 animals, 10 slices) by changing the frequency of stimulation for 2 min (horizontal bars) from 0.1 to 0.2 Hz (left) or 0.1 to 0.33 Hz (right). There are no significant differences ($P > 0.5$) in PPF or frequency facilitation between wild-type and mutant mice. **c**, Bath application of forskolin (10 μ M for 3 min) induced the same magnitude of potentiation in wild-type (white circles; 5 animals, 8 slices) and Rab3A-mutant mice (black circles; 5 animals, 7 slices).

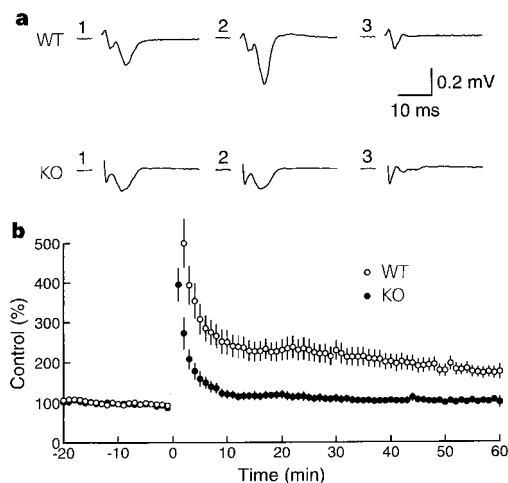


Figure 4 Mossy fibre LTP is blocked in Rab3A-deficient mice. **a**, Sample traces before (1), 55–60 min after tetanization (2) and after 10 μ M L-CCG1 bath application (3) in wild-type (WT, top) and knockout (KO, bottom) mice. **b**, LTP summary graphs (mean $\pm$ s.e.m.) in wild-type (white circles; 9 animals, 13 slices) and mutant mice (black circles; 10 animals, 13 slices). LTP was induced by one train lasting 5 s at 25 Hz in the presence of 100 μ M D-AP5.

average, the synapses in the two groups were activated to a similar extent by the tetanus.

Thus, in Rab3A-knockout mice, synaptic plasticity at the mossy fibre synapses exhibits a selective phenotype: LTP cannot be induced. All of the other synaptic parameters investigated were normal, suggesting that the phenotype is discrete. These data show that Rab3A is essential for mossy fibre LTP. To our knowledge, this observation provides the first example of a synaptic protein that is involved in vesicular transport and is essential for long-term synaptic plasticity, and supports the idea that mossy fibre LTP is a presynaptic modification operating at the secretory apparatus.

How might such a selective defect in mossy fibre LTP be explained? One model for mossy fibre LTP^{1,3} proposes that, during tetanic stimulation, Ca^{2+} accumulates in the presynaptic terminal and activates a Ca^{2+} /calmodulin-sensitive adenylyl cyclase. The resulting rise in cAMP activates PKA, which in turn phosphorylates substrate protein(s) which, in ways yet to be defined, increase the evoked release of transmitter. cAMP has also been shown to increase release at several other synapses^{28,29}. One possible explanation for the absence of mossy fibre LTP is that, in the Rab3A-deficient mouse, adenylyl cyclase is not activated because the tetanus-induced rise in presynaptic Ca^{2+} has been impaired. This seems unlikely, however, because three different Ca^{2+} -dependent forms of short-term plasticity at mossy fibre synapses (PPF, frequency facilitation and post-tetanic potentiation) were unaffected by the knockout. Another possibility is that Rab3A is involved in keeping adenylyl cyclase near the Ca^{2+} channels. In the absence of Rab3A, Ca^{2+} would no longer be able to activate the cyclase, but forskolin would still be effective in enhancing release because the cyclase is still present in the terminal. Although this hypothesis cannot formally be excluded, there is no evidence suggesting an interaction between Rab3A (or its associated proteins) and adenylyl cyclase.

A more likely possibility is that the Rab3A deletion prevents PKA-mediated modulation of some step in the release process itself. In this regard, the observation that Rab3A may regulate the number of vesicles that are released as a function of Ca^{2+} is important²⁴. This observation suggests a model by which LTP could occur, namely by changing the gain on the Ca^{2+} signal mediated by Rab3A. This hypothesis agrees well with the observation that both putative effectors for Rab3A, rabphilin and Rim, have PKA consensus

phosphorylation sites^{16,17}. According to this model, LTP would normally partially inactivate Rab3A, possibly by a phosphorylation-dependent action on its effectors, and no such inactivation would be possible in the absence of Rab3A. At this point, it is unclear whether Rab3A acts through rabphilin and/or Rim, or an as-yet unidentified effector. It is possible that rabphilin alone mediates mossy fibre LTP, and that the mislocation³⁰ and subsequent loss of rabphilin in the Rab3A-deficient mouse is the underlying cause for the mossy fibre LTP phenotype. Although the finding that PPF is unaltered at mossy fibre synapses but increased at CA1 excitatory synapses²⁴ might indicate a different role for Rab3A at these two types of synapse, the greater magnitude of PPF at mossy fibre synapses may mask the effects seen at CA1 synapses.

It is surprising that short-term synaptic plasticity and forskolin-induced enhancement are intact in the Rab3A-knockout mouse. This is unexpected because induction of mossy fibre LTP partly occludes PPF^{5,7}, frequency facilitation²⁶ and forskolin enhancement⁴. Although it cannot be ruled out that the involvement of Rab3A in mossy fibre LTP is independent of cAMP, the most plausible explanation for this apparent paradox is that induction of mossy fibre LTP normally also occludes short-term synaptic plasticity and forskolin enhancement by a mechanism that is parallel to, but not identical to, LTP induction. This explanation is supported by the observation that forskolin enhancement was also normal in mice that exhibited a loss of mossy fibre LTP as a result of mutations in the $\text{C}\beta_1$ and $\text{RI}\beta$ isoforms of PKA⁶.

In summary, we have found that knocking out Rab3A causes a remarkably selective functional defect at mossy fibre synapses: it abolishes mossy fibre LTP. This finding not only supports the presynaptic locus of mossy fibre LTP, but also provides the first link between a protein of the secretory machinery at the synapse and a form of long-term synaptic plasticity. Finally, our results are consistent with a model in which the increase in transmitter release during mossy fibre LTP occurs at a late step in the secretory process. □

Methods

Generation of the Rab3A-mutant mouse lines was described previously¹⁸. All data were acquired from homozygous mutant and wild-type littermates from heterozygote matings; heterozygotes were excluded. Hippocampal slices were prepared from mice 6 to 12 weeks old. The composition of the external recording solution was (in mM): 119 NaCl, 2.5 KCl, 1.2 MgSO_4 , 2.5 CaCl_2 , 1.0 NaH_2PO_4 , 26.2 NaHCO_3 and 10 glucose. Extracellular field recording and whole-cell patch-clamp recording² were performed at room temperature. For the whole-cell experiments, picrotoxin (100 μ M) was present to block inhibition, and 0.3 μ M CNQX and 50 μ M D-AP5 were added to prevent epileptiform discharge. Mice were shipped to San Francisco unidentified, and data were acquired and analysed by the experimenter in a blind fashion. After the experiment, the genotype was confirmed by PCR and western blot analysis using tail DNA and brain protein. Drugs used were D-2-amino-5-phosphonopivalic acid (D-AP5), (2S,3S,4S)-CCG/(2S,1'S,2'S)-2-(carboxycyclopropyl)-glycine (L-CCG1), 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX) (Tocris-Cookson), picrotoxin (Sigma) and forskolin (CalBiochem). The data are expressed as mean $\pm$ s.e.m. as a percentage of the baseline. Student's *t*-test was used to determine whether there was a significant difference in the means between the results from wild-type and mutant mice. For the anatomy presented in Fig. 1, sections of wild-type and Rab3A-deficient mice littermates (6 weeks old) were immunostained with the following antibodies: C142.2 (monoclonal antibody specific for Rab3A), C142.1 (monoclonal antibody specific for Rab3A/B/C), and a polyclonal antibody specific for synaptophysin II. The dilutions used were 1/2,000 to 1/5,000 of ascites or serum.

Received 28 February; accepted 13 June 1997.

- Nicoll, R. A. & Malenka, R. C. Contrasting properties of two forms of long-term potentiation in the hippocampus. *Nature* **377**, 115–118 (1995).
- Castillo, P. E., Weisskopf, M. G. & Nicoll, R. A. The role of Ca^{2+} channels in hippocampal mossy fiber synaptic transmission and long-term potentiation. *Neuron* **12**, 261–269 (1994).
- Tong, G., Malenka, R. C. & Nicoll, R. A. Long-term potentiation in cultures of single hippocampal granule cells: A presynaptic form of plasticity. *Neuron* **16**, 1147–1157 (1996).

4. Weisskopf, M. G., Castillo, P. E., Zalutsky, R. A. & Nicoll, R. A. Mediation of hippocampal mossy fiber long-term potentiation by cyclic AMP. *Science* **265**, 1878–1882 (1994).
5. Huang, Y.-Y., Li, X.-C. & Kandel, E. cAMP contributes to mossy fiber LTP by initiating both a covalently mediated early phase and macromolecular synthesis-dependent late phase. *Cell* **79**, 69–79 (1994).
6. Huang, Y. Y. *et al.* A genetic test of the effects of mutations in PKA on mossy fiber LTP and its relation to spatial and contextual learning. *Cell* **83**, 1211–1222 (1995).
7. Zalutsky, R. A. & Nicoll, R. A. Comparison of two forms of long-term potentiation in single hippocampal neurons. *Science* **248**, 1619–1624 (1990).
8. Weisskopf, M. G. & Nicoll, R. A. Presynaptic changes during mossy fiber LTP revealed by NMDA receptor-mediated synaptic responses. *Nature* **376**, 256–259 (1995).
9. López-García, J. C., Arancio, O., Kandel, E. R. & Baranes, D. A presynaptic locus for long-term potentiation of elementary synaptic transmission at mossy fiber synapses in culture. *Proc. Natl Acad. Sci. USA* **93**, 4712–4717 (1996).
10. Xiang, Z., Greenwood, A. C., Kairiss, E. W. & Brown, T. H. Quantal mechanism of long-term potentiation in hippocampal mossy-fiber synapses. *J. Neurophysiol.* **71**, 2552–2556 (1994).
11. Ferro-Novick, S. & Jahn, R. Vesicle fusion from yeast to man. *Nature* **370**, 191–193 (1994).
12. Bennett, M. K. & Scheller, R. H. A molecular description of synaptic vesicle membrane trafficking. *Annu. Rev. Biochem.* **63**, 63–100 (1994).
13. Südhof, T. C. The synaptic vesicle cycle: a cascade of protein–protein interactions. *Nature* **375**, 645–653 (1995).
14. DeCamilli, P., Benfenati, F., Valtorta, E. & Greengard, P. The synapsins. *Annu. Rev. Cell Biol.* **6**, 433–460 (1990).
15. Spillane, D. M., Rosahl, T. W., Südhof, T. C. & Malenka, R. C. Long-term potentiation in mice lacking synapsins. *Neuropharmacology* **34**, 1573–1579 (1995).
16. Fykse, E. M., Li, C. & Südhof, T. C. Phosphorylation of rabphilin-3A by Ca^{2+} /calmodulin- and cAMP-dependent protein kinases *in vitro*. *J. Neurosci.* **15**, 2385–2395 (1995).
17. Wang, Y., Okamoto, M., Schmitz, F., Hofmann, K. & Südhof, T. C. Rim is a putative Rab3 effector in regulating synaptic-vesicle fusion. *Nature* **387**, 000–000 (1997).
18. Geppert, M. *et al.* The role of Rab3A in neurotransmitter release. *Nature* **369**, 493–497 (1994).
19. Rosahl, T. W. *et al.* Essential functions of synapsins I and II in synaptic vesicle regulation. *Nature* **375**, 488–493 (1995).
20. Fykse, E. M. *et al.* Relative properties and localizations of synaptic vesicle protein isoforms: the case of the synaptophysins. *J. Neurosci.* **13**, 4997–5007 (1993).
21. Kamiya, H., Shinozaki, H. & Yamamoto, C. Activation of metabotropic glutamate receptor type 2/3 suppresses transmission at rat hippocampal mossy fibre synapses. *J. Physiol. (Lond.)* **493**, 447–455 (1996).
22. Yokoi, M. *et al.* Impairment of hippocampal mossy fiber LTD in mice lacking mGluR2. *Science* **273**, 645–647 (1996).
23. Zucker, R. S. Short-term synaptic plasticity. *Annu. Rev. Neurosci.* **12**, 13–31 (1989).
24. Geppert, M., Goda, Y., Stevens, C. F. & Südhof, T. C. Rab3A regulates a late step in synaptic vesicle fusion. *Nature* **387**, 810–814 (1997).
25. Regehr, W. G., Delaney, K. R. & Tank, D. W. The role of presynaptic calcium in short-term enhancement at the hippocampal mossy fiber synapse. *J. Neurosci.* **14**, 523–537 (1994).
26. Salin, P. A., Scanziani, M., Malenka, R. C. & Nicoll, R. A. Distinct short-term plasticity at two excitatory synapses in the hippocampus. *Proc. Natl Acad. Sci. USA* **93**, 13304–13309 (1996).
27. Seamon, K. B. & Daly, J. W. Forskolin: Its biological and chemical properties. *Adv. Cyclic Nucleot. Prot. Phosphoryl. Res.* **20**, 1–150 (1986).
28. Dixon, D. & Atwood, L. Adenylate cyclase system is essential for long-term facilitation at the crayfish neuromuscular junction. *J. Neurosci.* **9**, 4246–4252 (1989).
29. Schacher, S., Castellucci, V. F. & Kandel, E. R. cAMP evokes long-term facilitation in *Aplysia* sensory neurons that requires new protein synthesis. *Science* **240**, 1667–1669 (1988).
30. Li, C. *et al.* Synaptic targeting of rabphilin-3A, a synaptic vesicle Ca^{2+} /phospholipid-binding protein, depends on rab3A/3C. *Neuron* **13**, 885–898 (1994).

Acknowledgements. We thank H. Czerwonka for secretarial assistance. R.A.N. is a member of the Keck Center for Integrative Neuroscience and the Silvio Conte Center for Neuroscience Research. R.C.M. is a member of the Center for Neurobiology and Psychiatry, and the Center for the Neurobiology of Addiction. R.A.N. and R.C.M. are supported by grants from the NIH. T.C.S. was supported by grants from the HFJP and the W.M. Keck Foundation. R.J. was supported by a postdoctoral fellowship from DFG.

Correspondence and requests for materials should be addressed to R.A.N. (e-mail: nicoll@phy.ucsf.edu).

Rim is a putative Rab3 effector in regulating synaptic-vesicle fusion

Yun Wang*, Masaya Okamoto*, Frank Schmitz†, Kay Hofmann‡ & Thomas C. Südhof*

* Department of Molecular Genetics and Howard Hughes Medical Institute, The University of Texas, Southwestern Medical Center at Dallas, Texas 75235, USA

† Abteilung IX, Max-Planck-Institut für Experimentelle Medizin, 37075 Göttingen, Germany

‡ Swiss Institute for Experimental Cancer Research, 1066 Epalinges, Switzerland

Rab3 is a neuronal GTP-binding protein that regulates fusion of synaptic vesicles and is essential for long-term potentiation of hippocampal mossy fibre synapses^{1–5}. More than thirty Rab GTP-binding proteins are known to function in diverse membrane transport pathways, although their mechanisms of action are unclear. We have now identified a putative Rab3-effector protein

called Rim. Rim is composed of an amino-terminal zinc-finger motif and carboxy-terminal PDZ and C₂ domains. It binds only to GTP (but not to GDP)-complexed Rab3, and interacts with no other Rab protein tested. There is enrichment of Rab3 and Rim in neurons, where they have complementary distributions. Rab3 is found only on synaptic vesicles, whereas Rim is localized to presynaptic active zones in conventional synapses, and to presynaptic ribbons in ribbon synapses. Transfection of PC12 cells with the amino-terminal domains of Rim greatly enhances regulated exocytosis in a Rab3-dependent manner. We propose that Rim serves as a Rab3-dependent regulator of synaptic-vesicle fusion by forming a GTP-dependent complex between synaptic plasma membranes and docked synaptic vesicles.

Rab3A and Rab3C are small GTP-binding proteins of synaptic vesicles that cycle between vesicle-associated, GTP-complexed forms and cytosolic, GDP-complexed forms as a function of exocytosis^{6–8}. Deletion of Rab3A in mice enhances neurotransmitter release without changing the docking or priming of synaptic vesicles, suggesting that Rab3 has a role in regulating synaptic-vesicle fusion⁴. Rab3A is also required for long-term potentiation (LTP) of synaptic transmission in mossy fibre synapses of the hippocampus, indicating that Rab3 physiologically regulates the extent of synaptic-vesicle fusion in synaptic plasticity⁵. Rab3 and other Rab proteins are thought to function by GTP-dependent interactions with effector proteins^{8–10}, although the number and nature of these effectors are unknown: only a few candidate effectors have been described for more than thirty Rab proteins, including rabphilin-3A¹¹. Rabphilin is colocalized with Rab3 on synaptic vesicles, and binds to Rab3 as a function of GTP by means of an N-terminal zinc-finger domain^{12,13}. However, rabphilin on its own seems insufficient to account for the actions of Rab3, and so Rab3 could have multiple effectors.

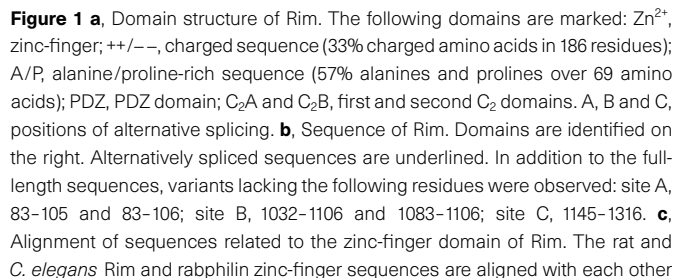
By using an activating form of Rab3C fused to the DNA-binding domain of LexA as a bait, we screened a rat brain cDNA library by yeast two-hybrid selection for proteins that interact with Rab3. Positive prey clones were rescued and retransformed into fresh yeast cells with Rab3 baits or control baits. We then evaluated the specificity of the prey and bait interactions by using two assays for transactivation, namely β -galactosidase induction and histidine autotrophy¹⁴. In this way we identified a class of prey clones that interact only with Rab3 in the yeast two-hybrid system, and encode part of a new protein called Rim (for Rab3-interacting molecule). Northern blots revealed that Rim mRNA is only detectable in brain (data not shown), suggesting that Rim is a candidate neuronal effector for the synaptic actions of Rab3.

We obtained only partial Rim cDNA clones in the yeast two-hybrid screens, so we isolated full-length cDNA clones from rat brain libraries. Their sequences showed that Rim is a large protein (1,553 amino acids) with several alternatively spliced variants (Fig. 1). We observed variable sequences in multiple independent cDNA clones at three positions. Two of these alternatively spliced sequences occur in three distinct variants, suggesting that Rim can be present in up to 18 forms. Databank searches identified a Rim homologue in *Caenorhabditis elegans* that is distributed over two cosmid sequences (T10A3 and K03A1). Because rabphilin is also present in *C. elegans*, both putative effectors for Rab3 are evolutionarily conserved. Analysis of the Rim sequence demonstrates that it is composed of multiple domains present in both the rat and *C. elegans* proteins: an N-terminal zinc-finger domain, a highly charged sequence followed by an alanine/proline-rich region, a PDZ domain, and C₂ domains separated by alternatively spliced sequences (Fig. 1).

The Rab3-binding protein rabphilin also has a zinc-finger motif and two C₂ domains. However, the zinc-fingers of Rim and rabphilin are only weakly homologous (42% identity over 55 residues), and their C₂ domains belong to different classes¹⁵. Furthermore, rabphilin lacks the other domains of Rim (the PDZ

We studied the interactions of the Rim prey and of a synaptotagmin control prey with wild-type and mutant forms of Rab3, Rab4, Rab5, Rab7, Rab17 and Rab22 by using yeast two-hybrid assays (Table 1). Most of the mutations locked the Rab proteins in either the GTP- or the GDP-bound states, and are referred to as activating and inactivating, respectively. We found that Rab3 bound to Rim, whereas the other Rab proteins were inactive (Table 1). Binding of wild-type Rab3A to Rim, as measured by β -galactosidase activation, was low. However, binding of Rab3A was greatly enhanced by independent activating mutations (such as Rab3A(Q81L) and Rab3A(S31V,S37P)). Corresponding mutations in other Rab proteins (such as Rab5(Q79L), Rab5(I53M) and Rab7(Q67L)) had no effect, even though the Rab5 mutants exhib-

Thus Rim binds exclusively to Rab3 as a function of GTP by means of a zinc-finger-containing sequence. We investigated whether similar sequences are present in other potential Rab-binding proteins by searching databanks with reiterative profiles constructed from the rat and *C. elegans* Rim and rabphilin zinc-finger sequences, and were able to identify a class of related zinc-



594

finger sequences with significant homology to the Rim/rabphilin zinc-finger (Fig. 1c). This class forms a separate subset of zinc-fingers that is present in proteins from yeast to mammals, some of which (Vac1, Fab1 and Hrs-2) function in membrane transport^{16–18}. Statistical analysis showed that the probability that the different zinc-fingers in Fig. 1c are not related to each other is $P < 0.06$. By contrast, no other zinc-finger domain reached scores of $P < 0.50$. Therefore, the zinc-finger domains of Rim and Rabphilin define a zinc-finger motif present in multiple proteins, some of which may function physiologically as Rab targets in membrane transport.

Table 1 Quantification of the interaction of Rim prey and Rab bait clones

Bait vector	pPreyRim-100	pPreyRim-52	pVP16-SytlC ₂ A/B
pLexNRab3A	37.2 ± 3.9	13.4 ± 2.5	9.7 ± 4.8
pLexNRab3A ^{T36N}	5.2 ± 2.8	5.6 ± 1.3	4.5 ± 2.3
pLexNRab3A ^{S31V/S37P}	513.2 ± 16.7	199.8 ± 27.0	3.2 ± 1.7
pLexNRab3A ^{Q81L}	601.6 ± 15.4	204.8 ± 20.4	5.0 ± 2.6
pLexNRab3A ^{N135I}	205.3 ± 12.4	55.6 ± 7.8	5.4 ± 2.7
pLexNRab3C*	583.3 ± 27.3	1,068.8 ± 16.6	4.1 ± 1.0
pLexNRab4B	6.7 ± 3.2	2.6 ± 0.7	2.9 ± 0.6
pLexNRab5	5.5 ± 3.1	4.3 ± 0.8	0.9 ± 0.5
pLexNRab5 ^{S34N}	6.4 ± 3.6	6.2 ± 1.3	4.2 ± 2.4
pLexNRab5 ^{I53M}	6.7 ± 3.3	7.2 ± 1.7	5.5 ± 2.9
pLexNRab5 ^{Q79L}	6.8 ± 2.3	6.4 ± 3.3	4.5 ± 3.0
pLexNRab7	4.1 ± 2.5	3.9 ± 0.6	6.5 ± 3.2
pLexNRab7 ^{Q67L}	5.9 ± 3.1	3.4 ± 0.7	8.4 ± 4.2
pLexNRab17	4.0 ± 2.7	3.7 ± 1.4	4.8 ± 2.8
pLexNRab22	1.2 ± 0.8	2.2 ± 1.2	1.3 ± 0.8

Results are obtained from yeast two-hybrid assays by transactivation of β -galactosidase activity. The two Rim prey vectors and the synaptotagmin 1 control prey vector (pVP16-SytlC₂A/B) were co-transformed with the indicated bait vectors into yeast and plated on –UTL plates. Single colonies were grown in liquid culture for 40 h before collection, and the β -galactosidase activity and protein content of each culture were determined. Data shown are mean ± s.e.m. nmol per min and mg protein from triplicate determinations. The sequence of rate pLexNRab3C* exhibits 7 amino-acid changes compared to the bovine sequence³⁰: Y17F, D24V, Q89R, V141A, V153I, H160R and N121S. These probably represent evolutionary changes, except for the Q89R mutation, which corresponds to the critical glutamine residue that is mutated in Rab-activating mutations.

The N-terminal zinc-finger domain of Rim is followed by a highly charged sequence containing phosphorylation consensus sequences for cyclic AMP- and Ca²⁺/calmodulin-dependent protein kinases, an alanine/proline-rich sequence, a PDZ domain, and two degenerate C₂ domains, which are separated by alternatively spliced sequences. In other proteins, PDZ domains recruit proteins to intercellular junctions. For example, the postsynaptic density protein PSD95 is thought to localize K⁺ channels, NMDA (N-methyl-D-aspartate) receptors and nitric oxide synthase to synaptic junctions, and ZO-1 and ZO-2 are thought to organize tight junctions (reviewed in ref. 19). The presence of a PDZ domain in Rim is surprising because no other known membrane transport protein contains a PDZ domain, and no protein described has both PDZ and C₂ domains. Rabphilin and synaptotagmins are membrane transport proteins that contain C₂ domains which bind Ca²⁺ by a defined consensus sequence¹⁵. In Rim, however, the C₂ domain consensus sequence and their Ca²⁺-binding sites are mutated. Furthermore, rabphilin and synaptotagmins have a conserved C-terminal sequence following the C₂ domains¹⁵ that is absent from Rim. Thus, although both Rim and rabphilin have C₂ domains, those of Rim occur in a different domain context and are structurally distinct, setting Rim apart from the vesicular C₂ domain proteins rabphilin and synaptotagmins.

Rabphilin levels are decreased by ~70% in Rab3A-knockout mice, probably because Rab3A recruits rabphilin to synaptic vesicles and mediates its transport from the cell body to synapses¹². To test whether Rim exhibits a similar dependence on Rab3A, we measured Rim levels in Rab3A-knockout mice. No change in Rim was observed in Rab3A-knockout mice in spite of a decrease in rabphilin and the similar Rab3-binding properties of the two proteins (data not shown). This result raises the possibility that Rim may not only be structurally different from rabphilin, but may also be subject to a different intracellular targeting pathway.

We used subcellular fractionation to address this question. Preliminary experiments suggested that Rim is enriched in synaptosomes and that it is particulate (data not shown). We therefore subfractionated synaptosomes into synaptic plasma membranes, synaptic vesicles, and mitochondria. Immunoblotting showed Rim to be highly enriched in synaptic plasma membranes, together with PSD95 and neuroligins (Fig. 2). Rim was absent from mitochondria, myelin and synaptic vesicles, whereas most of the Rab3A and rabphilin was associated with synaptic vesicles. Rim was also not detectable in highly purified synaptic vesicles, which did, however, contain abundant amounts of Rab3A and rabphilin (data not shown). In addition to the vesicle fraction, Rab3A and other vesicle proteins were also enriched in synaptic plasma membranes because they contain tightly docked synaptic vesicles. Taken together, these data suggest that Rim is coenriched with the PDZ-domain protein PSD95 in synaptic junctions but is absent from synaptic vesicles.

The biochemical results raise the possibility that Rim as a Rab3 target may be localized to the presynaptic plasma membrane, the target of synaptic vesicles that carry Rab3. This localization would agree well with the presence of a PDZ domain in Rim, a domain that is normally found in junctional plasma-membrane proteins¹⁹. We tested this hypothesis by immunocytochemistry of nerve terminals using immunofluorescence staining and silver-enhanced immunogold labelling. Immunofluorescence in spinal-cord motor neurons revealed a punctate staining pattern for Rim, with abundant labelling of synapses on the cell bodies of the motor neurons (Fig. 3A,a). Silver-enhanced immunogold labelling of the same synapses showed that Rim is exclusively presynaptic with no labelling outside nerve terminals. In the nerve terminals, we observed silver deposits close to the active zone but nowhere else; synaptic vesicles, the cytoplasm of the nerve terminals, and the plasma membrane outside the presynaptic active zone were not consistently labelled. (Fig. 3A,b).

Our data suggest that Rim may be specific for active zones in

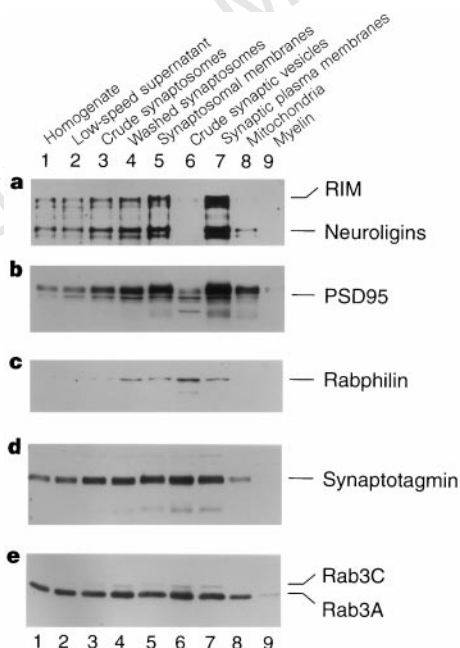


Figure 2 Localization of Rim to synaptic plasma membranes by subcellular fractionation. Rat brain homogenate (lane 1) was used to purify synaptosomes (lane 4), which were subfractionated into synaptic vesicles (lane 6), synaptic plasma membranes (lane 7), mitochondria (lane 8) and myelin (lane 9). Equivalent amounts of protein from each fraction were analysed by immunoblotting with antibodies to the proteins indicated in five independent blots: **a**, Rim/neuroligins; **b**, PSD95; **c**, rabphilin; **d**, synaptotagmin; **e**, Rab3A and Rab3C.

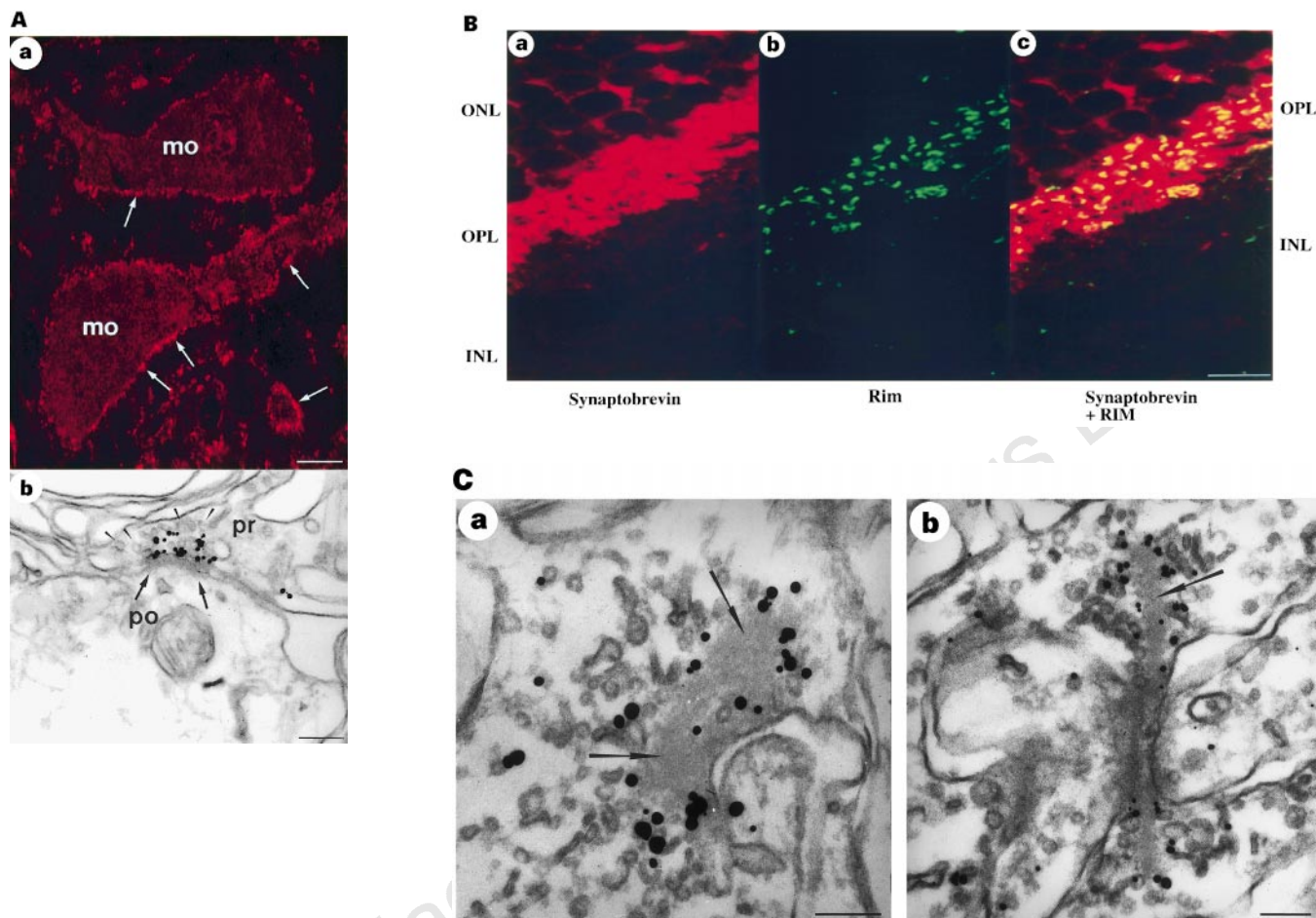


Figure 3 **A**, Localization of Rim in spinal cord motor neurons. **a**, Confocal micrograph of rat spinal-cord cryostat section stained with Rim primary and Cy2-conjugated secondary antibodies. Arrows point to synapses on motor-neuron cell bodies (mo). Scale bar, 50 μ m. **b**, Electron micrograph of a presynaptic terminal (pr) synapsing onto a postsynaptic neuron (po) stained for Rim by silver-enhanced immunogold labelling. Signals appear as black deposits found only in presynaptic nerve terminals close to the active zone. Scale bar, 0.2 μ m. **B**, Localization of synaptobrevin-2 and Rim in the outer plexiform layer of the retina by immunofluorescence microscopy. Confocal micrograph of a bovine retina cryostat section double-labelled with monoclonal antibodies to synaptobrevin-2

(a, red) and polyclonal antibodies to Rim (b, green). **c**, Images in **a** and **b** are merged, with areas positive for both antigens shown in yellow. Layers of the retina are identified on the left and right: ONL, outer nuclear layer; OPL, outer plexiform layer; INL, inner nuclear layer. Scale bar, 10 μ m. **C**, Immunoelectron microscopy of Rim in photoreceptor synapses. Electron micrograph of retinal synapses labelled for Rim by silver-enhanced immunogold staining. Ribbons (arrows) are shown in profile (**a**) or cross-section (**b**). Labelling is enriched on the surface of the ribbons but absent from synaptic vesicles and the plasma membrane outside the ribbons. Scale bar, 0.1 μ m.

nerve terminals. However, the small size of active zones in conventional synapses and the limited spatial resolution of silver-enhanced immunogold labelling make it difficult to determine whether the label is directly associated with the active zone. To resolve this, we studied ribbon synapses of retinal photoreceptor cells that contain synaptic ribbons instead of active zones. Ribbons are relatively large structures attached to the presynaptic plasma membrane, and can be easily visualized. Ribbons appear to dock synaptic vesicles and prepare them for exocytosis, in a manner analogous to the active zone of conventional synapses²⁰. Although ribbon synapses are structurally distinct from conventional synapses, they are probably similar mechanistically and contain the same synaptic proteins.

We performed double immunofluorescence labelling of retinæ with antibodies to Rim and to the synaptic-vesicle protein synaptobrevin II (Fig. 3b). Both proteins were highly enriched in synapses but were distributed in strikingly different patterns. Synaptobrevin was present throughout the synaptic layer because it is localized in all synaptic vesicles of the entire nerve terminal. In contrast, Rim was found only in a small area within the synapses. The pattern of Rim immunofluorescence in photoreceptor nerve terminals sug-

gests that its localization is restricted to ribbons. We confirmed this localization by immuno-electron microscopy and by biochemical isolation of ribbons. Silver-enhanced immunogold labelling revealed strong labelling of the surface of the entire synaptic ribbons (Fig. 3C). Synaptic vesicles or the adjacent plasma membranes, similar to conventional synapses, were largely unlabelled. Biochemical isolation of synaptic ribbons also demonstrated a high degree of enrichment of Rim with the detergent-insoluble ribbons (data not shown). Taken together, these data establish that Rim is present exclusively on ribbons in ribbon synapses and, to our knowledge, it is the first protein to be identified with such a localization.

By every indication, Rim should be involved in exocytosis. It is strategically localized to the active zone and to synaptic ribbons, and it specifically interacts with GTP-bound Rab3, which is localized to synaptic vesicles that dock to these structures. To confirm that Rim is involved in exocytosis, we investigated the effect of the N-terminal zinc-finger domain of Rim on the regulated secretion of growth hormone in transfected PC12 cells². When we co-transfected human growth hormone with mutant inactive tetanus toxin as a control into PC cells, we observed efficient Ca^{2+} -dependent growth-hormone

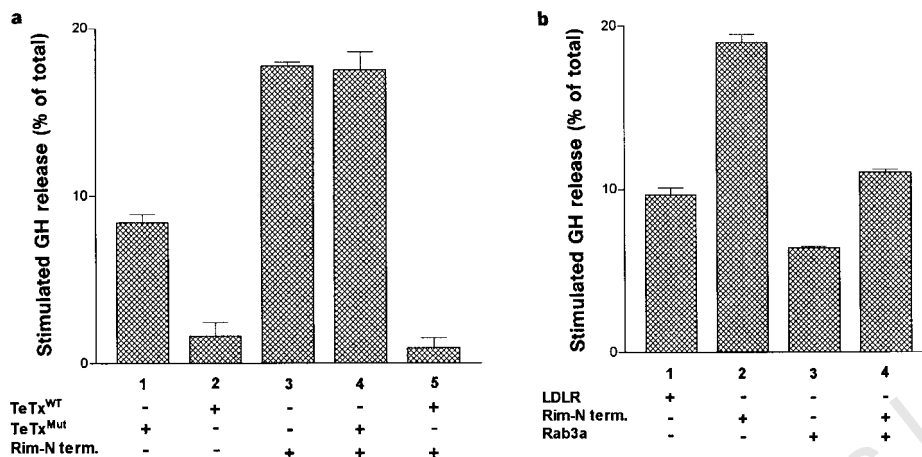


Figure 4 Rab3-dependent regulation of Ca^{2+} -dependent secretion in transfected PC12 cells. PC12 cells were co-transfected with plasmids encoding human growth hormone (GH) and the proteins indicated: TeTx^{WT}, light chain of wild-type tetanus toxin; TeTx^{Mut}, light chain of mutant inactive tetanus toxin²⁸; Rim-N term., N-terminal domains of Rim; LDLR, LDL receptors. Graphs show K^{+} -stimulated

release of growth hormone plotted as a percentage of total growth hormone expressed. Basal release was 6.5–8.5% of total, and did not change significantly as a function of transfection, except for wild-type tetanus toxin, which also decreased basal release. The increase in stimulated exocytosis induced by Rim remains sensitive to tetanus toxin and is reversed by Rab3A expression.

secretion triggered by membrane depolarization (Fig. 4a). In contrast, when we co-transfected wild-type tetanus toxin, Ca^{2+} -dependent secretion was largely suppressed. This confirms that the secretion observed is vesicular.

Co-transfection of the N-terminal domain of Rim with growth hormone, with or without the mutant tetanus toxin, resulted in a large increase in Ca^{2+} -dependent secretion that was not observed with any other co-transfected protein (Fig. 4a). The increase in secretion caused by Rim transfection was fully suppressed by wild-type tetanus toxin, showing that it is also due to vesicular exocytosis. Co-transfection of wild-type Rab4A with growth hormone, however, leads to a mild inhibition of triggered secretion² (Fig. 4b). The Rab3A inhibition and the Rim enhancement neutralize each other if both proteins are co-transfected. These data suggest that Rim has a functional role in secretion, although they do not establish whether Rim acts directly or indirectly. In agreement with the enhanced neurotransmitter release in the absence of Rab3A in hippocampal synapses⁴, our results indicate that interference with Rab3 function by expression of a truncated Rim protein promotes vesicular secretion.

Rab3A regulates synaptic-vesicle exocytosis by limiting the extent of Ca^{2+} -triggered membrane fusion⁴. This regulation is physiologically important for mossy fibre LTP, and possibly other forms of synaptic plasticity⁵. Rim is a candidate effector protein for Rab3 that may mediate the actions of Rab3 in membrane fusion. In Rim and rabphilin, a previously described putative Rab3 effector^{11,12}, Rab3 may have two distinct effector proteins. Although both of these effectors interact with Rab3 through related zinc-finger sequences, they are composed of different domains and localized to distinct subcellular compartments. Thus Rab3, and possibly other Rab proteins, could have multiple downstream functions mediated by diverse effectors. Rab3 and rabphilin are colocalized on synaptic vesicles^{12,13}. In contrast, Rim is only found on presynaptic active zones and synaptic ribbons, and is the first protein known to have this restricted localization. Other proteins of the synaptic plasma membrane, such as syntaxin and munc18-1, are distributed all over the plasma membrane, without enrichment in active zones. Because Rab3 is on synaptic vesicles, Rim can only bind to Rab3 when synaptic vesicles carrying GTP-Rab3 move to the active zone. As a consequence, the Rim–Rab3 complex forms only during or after docking of synaptic vesicles. The localization of Rim positions it for a role in synaptic-vesicle docking and fusion, in contrast to the

localization of rabphilin, which is present on free vesicles long before they reach the active zone¹². One possible model for Rab3 function is that the Rab3–Rim complex might serve as a clamp for Ca^{2+} -dependent exocytosis. Stochastic or regulated inactivation of this clamp by GTP hydrolysis or protein phosphorylation would then regulate the availability of vesicles for fusion, thereby determining how many vesicles fuse. Genetic and biochemical tests of this model are in progress. □

Methods

Molecular biology. Yeast-two hybrid bait vectors were constructed by standard techniques²¹ in pLexN using full-length Rab sequences with the indicated mutations. Yeast two-hybrid screens of a rat brain cDNA library in pVP16-3 were performed and evaluated as described^{14,22}. Two independent Rim prey clones encoding residues 1–345 (pPreyRim-100) and 11–399 (pPreyRim-52) were isolated. Liquid β -galactosidase assays were normalized for protein content²². A rat brain cDNA library in λ ZAPII was screened by standard methods²¹ using the Rim prey clone as a probe. Multiple overlapping clones covering the entire coding were sequenced to assemble the full-length sequence (GenBank accession number: AF007836). Rim–GST fusion protein vectors were obtained by cloning the *Eco*RI fragment from pPreyRim-52 into pGEX-KG23 (pGEXRIM52; N-terminal zinc-finger construct) or the 0.84-kilobase *Sma*I–*Pvu*II fragment into the *Sma*I site (pGEXRim-PDZ; this encodes residues 492–772).

Sequence analysis. Initial databank searches using BLAST software identified the domains in Rim (N-terminal zinc-finger, PDZ domain, and C₂ domains). The complete Rim *C. elegans* sequence was assembled from two cosmid sequences (T10A3 and K03A1; acc. nos U41035 and U41625, respectively). In order to define the relation of the Rim zinc-finger to other databank zinc-fingers, we applied the generalized profile method²⁴. A profile was constructed from a weighted multiple alignment of the zinc-finger regions from rat and *C. elegans* Rim and rabphilin using the BLOSUM45 substitution matrix²⁵ with a gap-creation penalty of 2.1 and gap-extension penalty of 0.2. Statistical significance of profile-matching scores was derived from the analysis of the score distribution of a locally shuffled database²⁶. Significant matches ($P < 0.08$) were incorporated into subsequent rounds of profile construction for iterative profile refinement. The initial profile search found significant matches to the zinc-finger regions of the nematode open reading frames C33D9.1 ($P < 0.06$) and ZK632.12 ($P < 0.08$). The two sequences were considered related to the Rim zinc-finger domain family and incorporated into the alignment for the next round of profile refinement. Three more profile refinement cycles, accepting only proteins with error probabilities of $P < 0.01$,

resulted in the protein family shown in Fig. 1c. The highest-scoring non-related sequences after those shown consisted of *trithorax* family members that reached error probabilities of $P > 0.5$ in all of the profile iterations.

Immunocytochemistry. Antibodies were raised in rabbits against the purified GST fusion proteins encoded by pGEX-Rim-52 and pGEX-Rim-PDZ. Double and single immunofluorescence labelling of cryostat sections from rat spinal cord and bovine retinae was performed²⁷ with two independent polyclonal Rim antibodies and multiple monoclonal antibodies to synaptic vesicle proteins. Staining was visualized by Cy2- and Cy3-conjugated secondary antibodies and viewed in a BioRad MRC1024 confocal microscope. Immuno-electron microscopy was performed by a pre-embedding protocol with silver enhancement²⁷. Controls for all immunocytochemistry experiments included the use of two independent antibodies, control stains with other antibodies, and experiments in which the first antibody was omitted.

PC12 cell transfections. PC12 cells (ATCC) were plated in collagen-coated 6-well dishes with 10^6 cells per well. Cells were transfected on day 1 with 0.2 μ g of Qiagen-purified pCMV5-hGH encoding human growth hormone and 1 μ g of the indicated expression plasmids using Lipofectamine (Life Technologies). Expression plasmids encoded the light chains of wild-type and mutant tetanus toxin in pCMV5 (ref. 28), LDL receptor and Rab3A in pCMV5, and residues 1–399 of Rim in pME18sf(–). On day 3, cells were collected, washed, and split into two portions, one of which was incubated for 20 min at 37 °C in control buffer (in mM) (145 NaCl, 5.6 KCl, 2.2 CaCl₂, 0.5 MgCl₂, 5.6 glucose, 0.5 ascorbate, 20 HEPES-NaOH, pH 7.4) and the other was incubated in the same buffer containing 56 KCl and 95 NaCl. After incubation, growth hormone in the medium and the cells was determined by radioimmunoassay (Nichols Institute). Secretion was calculated as the percentage growth hormone released as a function of stimulation. All experiments were done in triplicate at least three times.

Rim affinity chromatography. Glutathione–agarose columns without protein, with GST–Rim fusion protein encoded by pGEX-Rim-52, or with various control GST fusion proteins, were reacted with total rat brain homogenate prepared in 0.5% Triton X-100, 1 mM EDTA, 0.1 M NaCl, 0.1 g l⁻¹ PMSE, and 50 mM HEPES-NaOH, pH 7.4, with either 0.5 mM GDP- β S or GTP- γ S at 4 °C overnight. Samples were washed 5 times in the same buffer without nucleotides before analysis by SDS–PAGE and immunoblotting.

Other procedures. RNA blotting experiments were done using multiple tissue blots (Clontech). SDS–PAGE and immunoblotting were performed using standard procedures and antibodies described previously^{1,12,28,29}. Signals were detected by ECL and quantified by ¹²⁵I-labelled secondary antibodies. Protein assays were performed with the BioRad kit. Subcellular fractionations were performed as described²⁹ and verified by immunoblotting against a series of synaptic-vesicle, plasma-membrane and cytosolic antigens.

Received 19 February; accepted 22 May 1997.

- Geppert, M. *et al.* The role of Rab3A in neurotransmitter release. *Nature* **369**, 493–497 (1994).
- Holz, R. W., Brondyk, W. H., Senter, R. A., Kuizon, L. & Macara, I. G. Evidence for the involvement of Rab3A in Ca²⁺-dependent exocytosis from adrenal chromaffin cells. *J. Biol. Chem.* **269**, 10229–10234 (1994).
- Johannes, L. *et al.* The GTPase Rab3a negatively controls calcium-dependent exocytosis in neuroendocrine cells. *EMBO J.* **13**, 2029–2037 (1994).
- Geppert, M., Goda, Y., Stevens, C. F. & Südhof, T. C. Rab3A regulates a late step in synaptic vesicle fusion. *Nature* **387**, 810–814 (1997).
- Castillo, P. E., Janz, R., Südhof, T. C., Tzounopoulos, T., Malenka, R. C. & Nicoll, R. A. Rab3A is essential for mossy fibre long-term potentiation in the hippocampus. *Nature* **388**, 590–593 (1997).
- Burstein, E. S., Brondyk, W. H., Macara, I. G., Kaibuchi, K. & Takai, Y. Regulation of the GTPase cycle of the neuronally expressed Ras-like GTP-binding protein Rab3A. *J. Biol. Chem.* **268**, 22247–22250 (1993).
- Fischer von Mollard, G., Stahl, B., Khokhlatchev, A., Südhof, T. C. & Jahn, R. Rab3C is a synaptic vesicle protein that dissociates from synaptic vesicles after stimulation of exocytosis. *J. Biol. Chem.* **269**, 10971–10974 (1994).
- Südhof, T. C. Function of Rab3A GDP/GTP exchange. *Neuron* **18**, 519–522 (1997).
- Bourne, H. R., Sanders, D. A. & McCormick, F. The GTPase superfamily: a conserved switch for diverse cell functions. *Nature* **348**, 125–132 (1990).
- Nooffer, C. & Balch, W. E. GTPases: multifunctional molecular switches regulating vesicular traffic. *Annu. Rev. Biochem.* **63**, 949–990 (1994).
- Shirataki, H. *et al.* Rabphilin-3A, a putative target protein for smg p25A/rab3A p25 small GTP-binding protein related to synaptotagmin. *Mol. Cell. Biol.* **13**, 2061–2068 (1993).
- Li, C. *et al.* Synaptic targeting of rabphilin-3A, a synaptic vesicle Ca²⁺/phospholipid-binding protein, depends on rab3A/3C. *Neuron* **13**, 885–898 (1994).
- Stahl, B., Chou, J. H., Li, C., Südhof, T. C. & Jahn, R. Rab3 reversibly recruits rabphilin to synaptic vesicles by a mechanism analogous to raf recruitment by ras. *EMBO J.* **15**, 1799–1809 (1996).
- Vojtek, A. B., Hollenberg, S. M. & Cooper, J. A. Mammalian ras interacts directly with the serin/threonine kinase raf. *Cell* **74**, 205–214 (1993).
- Südhof, T. C. & Rizo, J. Synaptotagmins: C₂-domain proteins that regulate membrane traffic. *Neuron* **17**, 379–388 (1996).

- Weisman, L. S. & Wickner, W. Molecular characterization of VAC1, a gene required for vacuole inheritance and vacuole protein sorting. *J. Biol. Chem.* **267**, 618–623 (1992).
- Yamamoto, A. *et al.* Novel PI(4)P 5-kinase homologue, Fabp, essential for normal vacuole function and morphology in yeast. *Mol. Biol. Cell* **6**, 525–539 (1995).
- Bean, A. J., Seifert, R., Chen, Y. A., Sacks, R. & Scheller, R. H. Hrs-2 is an ATPase implicated in Ca²⁺-regulated secretion. *Nature* **385**, 826–829 (1997).
- Sheng, M. PDZs and receptor/channel clustering: rounding up the latest suspects. *Neuron* **17**, 575–578 (1996).
- Dowling, J. E. *The Retina. An Approachable Part of the Brain* (Belknap, Cambridge, MA, 1987).
- Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1989).
- Hata, Y. & Südhof, T. C. A novel ubiquitous form of munc18 interacts with multiple syntaxins. *J. Biol. Chem.* **270**, 13022–13028 (1991).
- Guan, K. L. & Dixon, J. E. Eukaryotic proteins expressed in *Escherichia coli*: an improved thrombin cleavage and purification procedure of fusion proteins with glutathione S-transferase. *Anal. Biochem.* **192**, 262–267 (1991).
- Bucher, P., Karplus, K., Moeri, N. & Hofmann, K. A flexible motif search technique based on generalized profiles. *Comput. Chem.* **20**, 3–23 (1996).
- Henikoff, S. & Henikoff, J. G. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA* **89**, 10915–10919 (1992).
- Hofmann, K. & Bucher, P. The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors. *Trends Biochem. Sci.* **20**, 347–349 (1995).
- Schmitz, F., Bechmann, M. & Drenckhahn, D. Purification of synaptic ribbons, structural components of the photoreceptor active zone complex. *J. Neurosci.* **15**, 7109–7116 (1996).
- McMahon, H. *et al.* Cellubrevin: a ubiquitous tetanus-toxin substrate homologous to a putative synaptic vesicle fusion protein. *Nature* **364**, 346–349 (1993).
- Ichtkenko, K. *et al.* Neurologin 1: A splice-site specific ligand for β -neurexins. *Cell* **81**, 435–443 (1995).
- Matsui, Y. *et al.* Nucleotide and deduced amino acid sequences of a GTP-binding protein family with molecular weights of 25,000 from bovine brain. *J. Biol. Chem.* **263**, 11071–11074 (1988).

Acknowledgements. We thank R. Janz, R. Jahn, M. Zerial and S. Butz for reagents, and J. L. Goldstein and M. S. Brown for advice and support. This work was partially supported by a grant from the HFSP.

Correspondence and requests for materials should be addressed to T.C.S. (e-mail: tsudho@mednet.swmed.edu).

Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II

Patrick D. Varga-Weisz, Matthias Wilm, Edgar Bonte, Katia Dumas, Matthias Mann & Peter B. Becker

European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany

Repressive chromatin structures need to be unravelled to allow DNA-binding proteins access to their target sequences. This de-repression constitutes an important point at which transcription and presumably other nuclear processes can be regulated^{1,2}. Energy-consuming enzyme complexes that facilitate the interaction of transcription factors with chromatin by modifying nucleosome structure are involved in this regulation^{3–5}. One such factor, nucleosome-remodelling factor (NURF), has been isolated from *Drosophila* embryo extracts^{4,6,7}. We have now identified a chromatin-accessibility complex (CHRAC) which uses energy to increase the general accessibility of DNA in chromatin. However, unlike other known chromatin remodelling factors, CHRAC can also function during chromatin assembly: it uses ATP to convert irregular chromatin into a regular array of nucleosomes with even spacing. CHRAC combines enzymes that modulate nucleosome structure and DNA topology. Using mass spectrometry, we identified two of the five CHRAC subunits as the ATPase ISWI, which is also part of NURF^{6,8}, and topoisomerase II. The presence of ISWI in different contexts suggests that chromatin remodelling machines have a modular nature and that ISWI has a central role in different chromatin remodelling reactions.

Chromatin reconstituted in the cell-free assembly system from *Drosophila* embryos^{9,10} is maintained in a dynamic state characterized by increased nucleosome mobility and overall DNA accessibility in the presence of ATP¹¹. Using ATP-dependent endonuclease cleavage of chromatin to assay for DNA accessibility¹¹, we purified and characterized a new type of chromatin remodelling

machine, CHRAC. CHRAC activity is sensitive to the detergent Sarkosyl (Fig. 1a). Addition of ATP to Sarkosyl-washed chromatin does not increase the accessibility of DNA (Fig. 1b, lanes 3–6); however, when CHRAC-containing fractions are added, DNA in chromatin becomes accessible to restriction enzymes (Fig. 1b, lane 7). Using this assay, we purified CHRAC from nuclear embryo extracts to apparent homogeneity (Fig. 2). On a gel filtration column, the purified activity ran as a complex with an apparent relative molecular mass of 670,000 (M_r 670K) (Fig. 2a). Consistently, proteins of 175K, 160K (which stains more strongly) and 130K, and two small proteins of M_r ~18K and 20K copurified with CHRAC activity with roughly equal stoichiometry (Fig. 2b).

CHRAC increased the accessibility of DNA in chromatin, but not of free DNA, in an energy-dependent manner. DNA in chromatin may be rendered accessible by, for example, alteration of the nucleosome core structure, loss of linker-DNA-binding proteins or of other chromatin-associated proteins, or by unfolding of the nucleosomal fibre. To test whether CHRAC activity required any factors other than nucleosomes, we reconstituted nucleosomes from pure histones on long linear DNA by salt-gradient dialysis. Addition of CHRAC and hydrolysable ATP was necessary and sufficient to increase the accessibility of the nucleosomal DNA to the restriction endonuclease *Dra*I (Fig. 3). As in the crude assembly system¹¹, only ATP and dATP supported CHRAC activity, whereas other nucleotides were inactive. This result identifies the nucleosome core particle as a target for CHRAC action.

The band corresponding to the 130K subunit was cut from a silver-stained SDS–polyacrylamide gel, cleaved in the gel by trypsin, and analysed by nano-electrospray mass spectrometry¹². Tandem

mass spectrometry of the tryptic cleavage products and computer-assisted sequence interpretation through database searches with peptide sequence tags¹³ led to the unambiguous identification of the 130K subunit as the known ATPase, ISWI^{6,8} (see Methods).

The 160K subunit was also analysed by mass spectrometry. Nine tryptic peptides were sequenced and the enzyme topoisomerase II (topo II; see Methods) was unambiguously identified. Both ISWI and topo II cofractionated precisely with CHRAC, as shown by western blotting, and topo II was active, as revealed by its ATP-dependent relaxation of negative supercoils (data not shown). By co-immunoprecipitation from crude CHRAC-containing fractions, we confirmed that topo II and ISWI were located in the same complex (Fig. 4). Topoisomerase II functions as a dimer¹⁴, which is consistent with the staining of the 160K topo II band in CHRAC being stronger than the other subunits (Fig. 2). Assuming that two molecules of topo II are present, the relative molecular masses of the individual subunits add up to the observed M_r of 670K.

ISWI is also a subunit of NURF, another nucleosome-remodelling complex that facilitates the binding of transcription factors to chromatin. No other subunit is shared by the complexes^{4,6}. Two hSNF2L(the human ISWI homologue)-containing complexes of ~700K and 450K have been identified¹⁵, corresponding to the M_r determined here for CHRAC and with that of NURF⁴, suggesting the existence of human counterparts for both ISWI-containing complexes.

The ATPase activity of NURF is stimulated by nucleosomal but not free DNA⁶. By contrast, the ATPase activity of topo II is stimulated by free DNA¹⁶. Comparison of the ATPase activities of CHRAC with that of purified *Drosophila* topoisomerase II (Fig. 5)

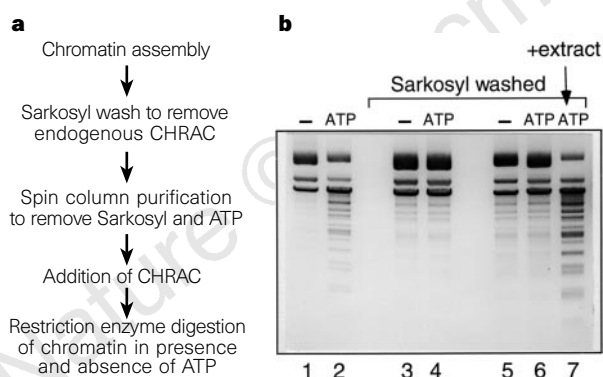


Figure 1 CHRAC assay. **a**, Scheme of the assay. **b**, Lanes: 1, 2, endogenous CHRAC activity revealed by enhanced *Dna*I cleavage of reconstituted chromatin in the presence of ATP; 3–6, Sarkosyl-stripped chromatin becomes more accessible in general, but lacks the ATP-dependent opening; 7, addition of 1 µl CHRAC-containing extract to stripped chromatin reconstitutes ATP-dependent endonuclease accessibility. Chromatin in lanes 3, 4 was washed with 0.075% Sarkosyl, chromatin in lanes 5–7 with 0.15% Sarkosyl. A negative image of the ethidium bromide-stained agarose gel is shown.

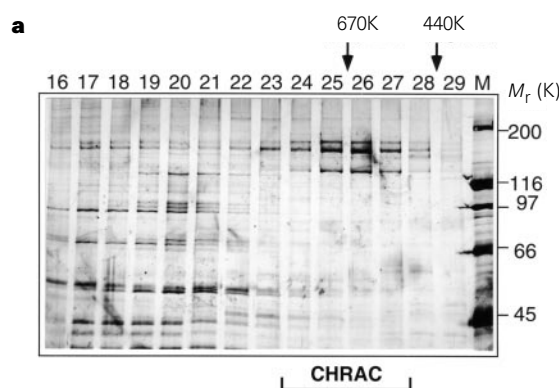


Figure 2 Purification of CHRAC. **a**, 8% SDS–PAGE of fractions from the Superose-6 gel filtration column, calibrated with thyroglobulin (M_r 670K) and ferritin (440K) size standards (Pharmacia). CHRAC activity ("CHRAC") peaks in fractions 25 and 26. **b**, Purified CHRAC (CHR) was resolved by 8% and 15% SDS–PAGE and

stained with silver. M: size markers (BioRad) with indicated relative molecular masses in thousands. Asterisks: small proteins that co-fractionate with CHRAC activity. These proteins run with histone markers on other gels.

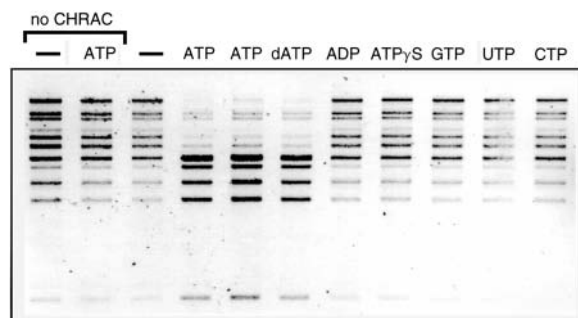


Figure 3 CHRAC targets nucleosomes. Nucleosomes were reconstituted on linear DNA from *Drosophila* histones by salt-gradient dialysis. 28 ng of fraction 25 (Fig. 2) or buffer (no CHRAC) were added to ~80 ng DNA reconstituted into chromatin in EX120 in the presence of 200 $\mu\text{g ml}^{-1}$ BSA and the indicated nucleotides at 1 mM. After *DraI* digestion, the purified DNA was separated on an agarose gel. A negative image of the ethidium bromide-stained gel is shown.

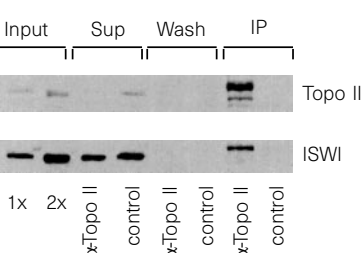
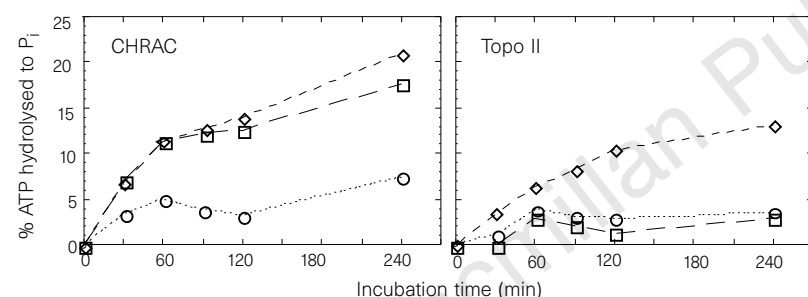


Figure 4 ISWI and topo II co-immunoprecipitate from the crude first CM-Sepharose eluate. Topo II was immunoprecipitated with a protein A-Sepharose-purified rabbit polyclonal antibody (α -topoII) or a nonspecific control antibody. 10% (1x) and 20% (2x) of the input material, 10% of the supernatant (sup), 20% of the final wash and the entire immunoprecipitate (IP) was subjected to consecutive western analysis with an antibody against ISWI and topo II.

Figure 5 ATPase activities of purified CHRAC and topoisomerase II. The ATPase activity of roughly equivalent amounts (with respect to topoisomerase activity) of CHRAC (left) and purified topoisomerase II (right) were assayed in the presence of control buffer (circles), free DNA (diamonds) and nucleosomes (squares). The percentage of free phosphate was determined at the indicated reaction times.

showed that topoisomerase II was only stimulated by free DNA as well as by nucleosomes, whereas CHRAC was stimulated by free DNA as well as by nucleosomes. We assume that the nucleosome-stimulated ATPase activity in CHRAC is due to ISWI, which is common to both CHRAC and NURF.

CHRAC and NURF increase the accessibility of chromatin by different mechanisms. CHRAC and NURF preparations, standardized by their content of immunoreactive ISWI, had different activities when compared in accessibility assays: CHRAC was less effective in the GAGA-factor-dependent nucleosome remodelling assay than NURF; conversely, NURF did not stimulate the sensitivity to restriction endonucleases (data not shown). NURF perturbs histone–DNA interactions in mononucleosomes when added at a stoichiometry of one NURF per 20 nucleosome core particles⁴. The SWI/SNF and RSC complexes also perturb histone–DNA interactions, but require higher stoichiometries^{1,5}. So far, we have not observed a similar effect when incubating equimolar amounts of CHRAC and mononucleosomes under comparable experimental conditions (unpublished observations).

The most striking difference between CHRAC and NURF, however, was in their effect on nucleosomal arrays. The assembly of nucleosomes is rapid in the *Drosophila* and related *Xenopus* systems in the absence of ATP, but correct nucleosome spacing requires time and energy^{9,17}. The ATP-dependent spacing principle has so far remained elusive. Addition of CHRAC to chromatin assembled in the absence of ATP, and therefore lacking regular nucleosome spacing, resulted in an ATP-dependent alignment of nucleosomes which became evident from the new regularity of the micrococcal nuclease cleavage pattern (Fig. 6a). This result identifies CHRAC as an ATP-dependent nucleosome-spacing factor. NURF has the opposite effect: addition of large amounts of NURF to a nucleoso-

mal array perturbs the regular repeat pattern⁴. The side-by-side comparison of NURF and CHRAC (standardized by their ISWI content) in our nucleosome-spacing assay demonstrated the difference between the two activities (Fig. 6b). Irregular chromatin obtained by only brief chromatin assembly was converted into a regular nucleosomal array in an ATP-dependent reaction by CHRAC, but not by NURF.

We propose two non-exclusive mechanisms by which CHRAC may mediate these two apparently opposing effects of increasing the accessibility of nucleosomal DNA and of imposing regularity on nucleosomal arrays. CHRAC might increase the mobility of nucleosomes with the expenditure of energy, a reaction that occurs in a cell-free system¹¹. Enhanced nucleosome mobility may increase the chance of transient exposure of a particular sequence in accessible linker regions and at the same time facilitate the alignment of nucleosomes into a regular repeat structure. Alternatively, CHRAC may participate in nucleosome assembly and disassembly, like the histone chaperones which, depending on the circumstance, may function either as histone donors in nucleosome assembly or as histone acceptors during disassembly of nucleosome cores¹⁸. The identification of the ATPase ISWI in two structurally and functionally distinct chromatin remodelling complexes highlights the modular nature of these complexes and suggests that ISWI may drive diverse nucleosome remodelling reactions specified by different interaction partners. ISWI, being found in both NURF and CHRAC, may be directly involved in targeting and rearranging nucleosome structure.

The association of ISWI and topo II in one complex is interesting. The topoisomerase activity (DNA-strand breakage and religation) does not seem to be involved in CHRAC function: VP16, a specific inhibitor of topo II (ref. 19), did not prevent CHRAC from

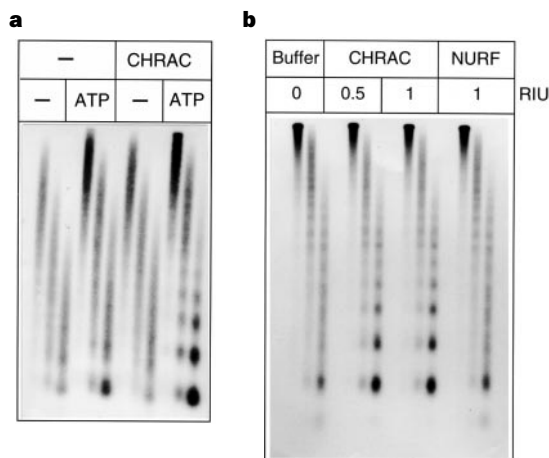


Figure 6 **a**, CHRAC is an ATP-dependent spacing factor. Chromatin was assembled without ATP, Sarkosyl-stripped to remove endogenous CHRAC, and the effect of addition of ATP and/or CHRAC on the regularity of the nucleosomal array was studied by micrococcal nuclease analysis as described in Methods. **b**, CHRAC and NURF are functionally distinct. Equivalent amounts of CHRAC and NURF (expressed as relative ISWI units, RIU) were compared for their ability to create a regular nucleosomal array.

increasing chromatin accessibility to restriction enzymes, and only slightly altered its effect on nucleosome spacing (data not shown). Purified topo II could not substitute for CHRAC in either of these reactions. The physiological importance of the association of ISWI and topoisomerase II in one complex will only become apparent once further CHRAC subunits have been identified. Nucleosome remodelling by ISWI may help the interaction of topo II with chromatin and thus enhance its function. Conversely, the association of topo II with chromosomes may localize CHRAC activity to specific chromosomal sites. Topoisomerase II is involved in many cellular processes such as the removal of catenates after replication, chromosome condensation and decondensation, kinetochore assembly and sister chromatid separation at mitosis (for reviews, see refs 16, 20–22); it is a major structural component of the mitotic chromosome axis^{23,24} and is distributed widely over interphase chromosomes, but is also concentrated at specific heterochromatic sites²⁵. Distinct populations of topo II with dynamic nuclear localization patterns during the cell cycle have been visualized²⁵. The existence of topo II isoforms, cell-cycle-regulated post-translational modifications, and various associated proteins contribute to the structural and functional diversity of topo II¹⁶. The CHRAC-associated topo II represents only a fraction of the total topo II in a nucleus and may therefore correspond to a subpopulation dedicated to a specific function. We recently observed a stimulatory effect by CHRAC in a cell-free replication system: CHRAC, but not purified topo II, enhanced the initiation of replication from a nucleosomal SV40 origin (V. Alexiadis, P.D.V., P.B.B. and C. Gruss, manuscript in preparation).

Methods

Nucleosome reconstitution and CHRAC assay. Chromatin was assembled in *Drosophila* embryo extracts as described^{9–11}. After the assembly, 0.075% Sarkosyl (sodium lauroylsarcosine) was added to disrupt endogenous CHRAC activity. Sarkosyl and ATP were removed by gel filtration¹¹. CHRAC assay: 10 μ l Sarkosyl-washed chromatin (equivalent to ~40 ng DNA), 0.5–5 μ l CHRAC-containing fractions, 10–50 units of *Dra*I or *Hinc*II and 1–2 mM ATP in 40 μ l EX-120 buffer¹¹ were incubated for 30–60 min at 26°C.

Nucleosome reconstitution by salt-gradient dialysis. 4 μ g λ DNA, 2 μ g *Drosophila* histones and 50 μ g BSA (Fermentas) were mixed in 100 μ l HSB

(10 mM HEPES-KOH, pH 7.6, 1 mM EDTA, 1 mM 2-mercaptoethanol, 2 M NaCl) buffer. This mixture was dialysed in a microcollodion bag (Sartorius) against 600 ml HSB. During dialysis over 2 d at 4°C, the salt was gradually removed by exchange of dialysis buffer with buffer lacking NaCl but containing 0.05% NP-40.

CHRAC purification. A comprehensive protocol is available upon request. CHRAC was purified from nuclear extracts of 12–16-h-old *Drosophila* embryos. Nuclei were extracted with ammonium sulphate²⁶ and extracts were dialysed into CB-150 buffer (CB: 10 mM HEPES-KOH, pH 7.6, 1 mM MgCl₂, 0.5 mM EGTA, 0.1 mM EDTA, 0.5 mM sodium metabisulphite, 0.2 mM PMSF, 1 mM DTT; CB-150 contains 150 mM KCl) before precipitation with 20% methanol. All columns were equilibrated in CB-150 and loaded with fractions diluted to a conductivity equal to CB-150. Proteins were chromatographed on a CM-Sephacrose FF column (all columns and resins from Pharmacia) and CHRAC activity was eluted with CB-400. Active fractions were rechromatographed on a smaller CM-Sephacrose FF column from which the activity was eluted with CB-200. From the following Mono-Q HR 10/10 column, CHRAC eluted at a conductivity equal to CB-350. Successive concentrations involved a 1-ml CM-Sephacrose FF column and a Filtron Microsepe concentrator (*M_r* cutoff, 30K). The material was separated on a Superose-6 10/30 column in 20 mM HEPES-KOH, pH 7.6, 400 mM KCl, 1 mM MgCl₂, 0.5 mM EGTA, 1 mM DTT, 20% glycerol. If necessary, a final fractionation on a Mono-Q 1/1 column was added from which CHRAC was eluted by a salt gradient.

Estimations of purification factor and yield are complicated by the presence of CHRAC inhibitors in crude nuclear extracts. As the purification after the CM-Sephacrose step leads to a 200–400 fold enrichment of CHRAC activity, we estimate the overall degree of purification to be at least 1,000-fold. Unless stated otherwise, the CHRAC concentration in the experiments was 250 μ g ml^{–1}, as determined by a modified Bradford assay (Biorad) with BSA as standard.

Mass spectrometry. Details of the analysis are available upon request. Proteins were proteolytically trypsinized in the gel^{12,27}. Extracted peptides were purified on a 100 nl R2 Poros microcolumn and step-eluted in 1 μ l 60% methanol/5% formic acid into a nanoelectrospray capillary. Mass spectrometry was done on an API III triple quadrupole instrument (PE-Sciex, Ontario) equipped with an upgraded collision cell and the nanoelectrospray ion source²⁸. For the 130K subunit, the sequence tag (776.6)(I/L)EQ(1146.4) from the fragment spectrum of a peptide (*M_r* 1,877) identified uniquely the tryptic peptide comprising amino acids 736–748 of the ISWI protein in a non-redundant database (data not shown). Four other ISWI peptides were sequenced from the same digest corresponding to amino acids 160–171, 394–402, 450–458 and 638–648. From the tryptic peptide mixture of the 160K protein, 9 peptides (T1–T9) were sequenced which identified unambiguously *Drosophila* topoisomerase II (SwissProt P15348; T₁: ITFSPDLAK, T₂: DLYGVFPLR, T₃: SLAVSGLGVIGR, T₄: SGIVESVLAWAK, T₅: YIFTIMSPLTR, T₆: IFDEILVNAADNK, T₇: ELMWVVDNSQNR, T₈: SQLEHILLRPDSYIGSV EFTK, T₉: GFQQVSFVNSIATYK; data not shown).

ATPase assay. Nucleosomes were reconstituted on a 146-bp DNA fragment by salt-gradient dialysis at high histone input to ensure that no free DNA contaminated the preparation. The reaction (9 μ l) contained: 60 ng DNA or the equivalent amount of mononucleosomes, 6.6 mM HEPES, pH 7.6, 0.66 mM EDTA, 0.66 mM 2-mercaptoethanol, 0.033% N-P40, 1.1 mM MgCl₂, 33 μ M ATP, 5 μ Ci [γ -³²P]ATP (3,000 Ci mmol^{–1}, Amersham) and 14 ng CHRAC or 11.5 units topo II (USB) or CHRAC400 buffer. Unreacted ATP and free phosphate were separated by thin-layer chromatography and quantified by using PhosphorImager and Imagequant software.

Co-immunoprecipitation. Equal amounts of protein A-Sephacrose-purified rabbit polyclonal antibody against topoisomerase II (W4-250-1, provided by D. Arndt-Jovin) or a control protein A-purified rabbit IgG were added to 10 μ l of the first CM-Sephacrose eluate. 5 μ l IP-150 (20 mM HEPES, pH 7.6, 0.5 mM EGTA, 10% glycerol, protease inhibitor, 150 mM KCl, 5 mM MgCl₂), 1 μ l (20 μ g) chicken ovalbumin and 2 μ l of 3 M KCl were added. This mixture was kept for 1 h on ice, added to protein A-Sephacrose beads (10 μ l settled volume) in 50 μ l IP-400 (as IP-150, but containing 400 mM KCl, 1 mM MgCl₂, 50 μ g ovalbumin and 0.1% N-P40) and incubated further for 2 h at 4°C. Beads were washed twice with 50 μ l IP-400, once with IP-150, and resuspended in 25 μ l IP-150. 5 μ l 6 $\times$ SDS loading buffer were added and the sample was

analysed by SDS-PAGE and western blotting. Rabbit anti-ISWI antibody was biotinylated using the Clontech Biotin-X-NHS ester labelling kit. Detection was by streptavidin-peroxidase polymer (Sigma) with an ECL detection kit (Amersham).

Nucleosome spacing. *Drosophila* embryo assembly extract was dialysed against EX120, 0.2 mM PMSF, 0.5 mM sodium metabisulphite to remove ATP. 60 µl of extract, 66 µl EX50, 14 µl 30 mM MgCl₂, 10 mM DTT, 600 mM NaCl and 1 µg plasmid DNA were mixed and incubated for 3 h at 26 °C. Chromatin was Sarkosyl-stripped and purified by gel filtration^{7,11}. 20 µl of this chromatin (~70 ng DNA) plus 30 µl EX120 were incubated for 90 min with either 1 µl CHRAC or buffer in presence or absence of 2 mM ATP. The nucleosome repeat was revealed with an oligonucleotide hybridizing to a GAGA element of the *hsp26* promoter as described^{17,29}. 0.1 and 0.2 µl CHRAC and 1 µl NURFP11 (gift from T. Tsukiyama and C. Wu) were assayed in the experiment shown in Fig. 6b. The CHRAC preparation contained 5× more ISWI protein than the NURF preparation.

Received 29 January; accepted 28 May 1997.

- Kingston, R., Bunker, C. & Imbalzano, A. N. Repression and activation by multiprotein complexes that alter chromatin structure. *Genes Dev.* **10**, 905–920 (1996).
- Felsenfeld, G. Chromatin unfolds. *Cell* **86**, 13–19 (1996).
- Peterson, C. L. & Tamkun, J. W. The SWI-SNF complex: a chromatin remodelling machine? *Trends Biochem. Sci.* **20**, 143–146 (1995).
- Tsukiyama, T. & Wu, C. Purification and properties of an ATP-dependent nucleosome remodelling factor. *Cell* **83**, 1011–1020 (1995).
- Cairns, B. R. *et al.* RSC, an essential, abundant chromatin-remodelling complex. *Cell* **87**, 1249–1260 (1996).
- Tsukiyama, T., Daniel, C., Tamkun, J. & Wu, C. ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kD subunit of the nucleosome remodelling factor. *Cell* **83**, 1021–1026 (1995).
- Tsukiyama, T., Becker, P. B. & Wu, C. ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. *Nature* **367**, 525–532 (1994).
- Elfring, L. K., Deuring, R., McCallum, C. M., Peterson, C. L. & Tamkun, J. W. Identification and characterization of *Drosophila* relatives of the yeast transcriptional activator SNF2/SWI2. *Mol. Cell. Biol.* **14**, 2225–2234 (1994).
- Becker, P. B. & Wu, C. Cell-free system for assembly of transcriptionally repressed chromatin from *Drosophila* embryos. *Mol. Cell. Biol.* **12**, 2241–2249 (1992).
- Becker, P. B., Tsukiyama, T. & Wu, C. Chromatin assembly extracts from *Drosophila* embryos. *Meth. Cell Biol.* **44**, 207–223 (1994).
- Varga-Weisz, P. D., Blank, T. A. & Becker, P. B. Energy-dependent chromatin accessibility and nucleosome mobility in a cell-free system. *EMBO J.* **14**, 2209–2216 (1995).
- Wilm, M. *et al.* Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* **379**, 466–469 (1996).
- Mann, M. & Wilm, M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Analyt. Chem.* **66**, 4390–4399 (1994).
- Berger, J. M., Gambin, S. J., Harrison, S. C. & Wang, J. C. Structure and mechanism of DNA topoisomerase II. *Nature* **379**, 225–232 (1996).
- Wang, W. D. *et al.* Diversity and specialization of mammalian SWI/SNF complexes. *Genes Dev.* **10**, 2117–2130 (1996).
- Wang, J. C. DNA topoisomerases. *Annu. Rev. Biochem.* **65**, 635–692 (1996).
- Almouzni, G. & Méchali, M. Assembly of spaced chromatin. Involvement of ATP and DNA topoisomerase activity. *EMBO J.* **7**, 4355–4365 (1988).
- Walter, P., Owen-Hughes, T. A., Coté, J. & Workman, J. L. Stimulation of transcription factor binding and histone displacement by nucleosome assembly protein 1 and nucleoplasmin requires disruption of the histone octamer. *Mol. Cell. Biol.* **15**, 6178–6187 (1995).
- Chen, G. L. *et al.* Nonintercalative antitumor drugs interfere with the breakage-reunion reaction of mammalian DNA topoisomerase II. *J. Biol. Chem.* **259**, 13560–13566 (1984).
- Buchanan, P., Saumweber, H. & Arndt, J. D. Consequences of topoisomerase II inhibition in early embryogenesis of *Drosophila* revealed by *in vivo* confocal laser scanning microscopy. *J. Cell Sci.* **104**, 1175–1185 (1993).
- Poljak, L. & Käs, E. Resolving the role of topoisomerase II in chromatin structure and function. *Trends Cell Biol.* **5**, 348–354 (1995).
- Wartburton, P. E. & Earnshaw, W. C. Untangling the role of DNA topoisomerase II in mitotic chromosome structure and function. *BioAssays* **19**, 97–99 (1997).
- Earnshaw, W. C., Halligan, B., Cooke, C. A., Heck, M. M. & Liu, L. F. Topoisomerase II is a structural component of mitotic chromosome scaffolds. *J. Cell Biol.* **100**, 1706–1715 (1985).
- Gasser, S. M., Laroche, T., Falquet, J., Boy de la Tour, E. & Laemmli, U. K. Metaphase chromosome structure. Involvement of topoisomerase II. *J. Mol. Biol.* **188**, 613–629 (1986).
- Swedlow, J. R., Sedat, J. W. & Agard, D. A. Multiple chromosomal populations of topoisomerase II detected *in vivo* by time-lapse, three-dimensional wide-field microscopy. *Cell* **73**, 97–108 (1993).
- Sandaltzopoulos, R., Mitchelmore, C., Bonte, E., Wall, G. & Becker, P. B. Dual regulation of the *Drosophila hsp26* promoter *in vitro*. *Nucleic Acids Res.* **23**, 2479–2487 (1995).
- Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Analyt. Chem.* **68**, 850–858 (1996).
- Wilm, M. & Mann, M. Analytical properties of the nano-electrospray ion source. *Analyt. Chem.* **68**, 1–8 (1996).
- Wall, G., Varga-Weisz, P. D., Sandaltzopoulos, R. & Becker, P. B. Chromatin remodelling by GAGA factor and heat shock factor at the hypersensitive *Drosophila hsp26* promoter *in vitro*. *EMBO J.* **14**, 1727–1736 (1995).

Acknowledgements. This research was supported by an EMBO fellowship to P.D.V. and by a grant from the DFG to P.B.B. We thank T. Tsukiyama and C. Wu for a generous gift of NURF fractions and ISWI antibodies, S. Beek and J. Tamkun for antibodies against ISWI, P. Fisher and D. Arndt-Jovin for antibodies against topo II; A. Shevchenko for technical help, H. Stunnenberg for important suggestions and discussions, T. Blank and C. Garcia-Jimenez for introduction to chromatography and P. Riedinger for artwork.

Correspondence and requests for materials should be addressed to P.B.B. (e-mail: becker@embl-heidelberg.de).

erratum

Modelling teleconnections between the North Atlantic and North Pacific during the Younger Dryas

Uwe Mikolajewicz, Thomas J. Crowley, Andreas Schiller & Reinhard Voss

Nature **387**, 384–387 (1997)

In Figs 1 and 3b, the numbering was very faint; also, shading was lost from Fig. 4. The corrected figures 1, 3b and 4 are shown here. □

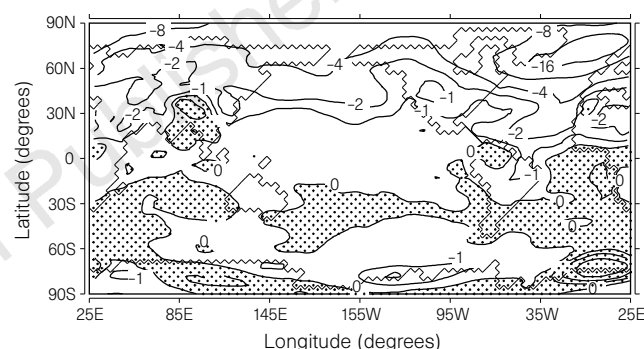


Figure 1

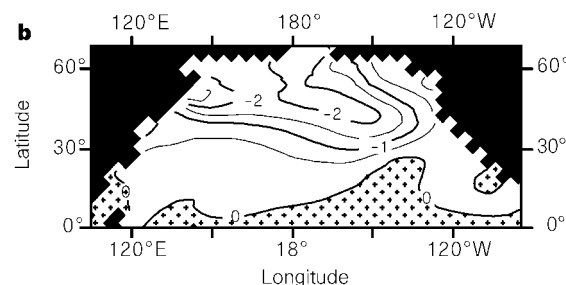


Figure 3

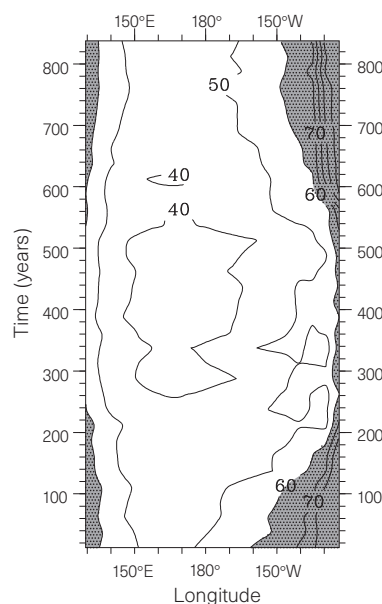


Figure 4

The breakdown on bioinformatics

Hardware and software make up this survey – items include software for genomics work, DNA and protein sequence analysis, protein viewing, and a number of products for drug discovery and drug screening applications.

Bioinformatics Workbench

From MDL Information Systems

An intranet-based 'virtual laboratory' that delivers data and computing tools for genomic research

Developed at the National Center for Supercomputing Applications (NCSA) at the University of Illinois, the Bioinformatics Workbench provides an environment that integrates the genome-related tools and databases emanating from sequencing efforts on many plant and animal species. This system is designed to tie together software that is already available. Workbench offers a platform-independent interface to more than 100 public domain databases and software packages for DNA and protein sequence analysis. With this software, scientists connect via a Web browser on their corporate intranets to search all major molecular and structural biology databases. The Bioinformatics Workbench uses a generic design that allows tools, databases, servers and computer engines based on different formats to interoperate. The interface to these components works with any standard forms-capable, HTML browser. Previously known as Biology Workbench, the program was developed by the Computational Biology Group of the NCSA. The NCSA version is available free to the academic community at <<http://bioweb.ncsa.uiuc.edu>>

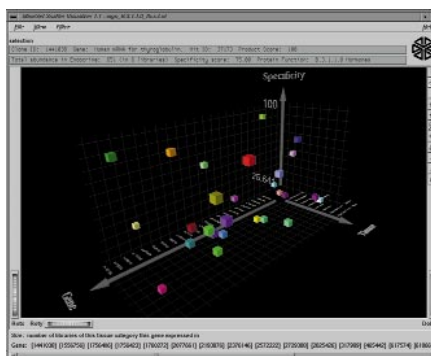
Reader Enquiry No. 100

LifeTools 3D

From Incyte Pharmaceuticals

Data mining and visualization software for searching genomic information

The program displays genomic data as interactive, multidimensional graphics. It can present up to five dimensions of data simultaneously, allowing scientists to collate factors, such as gene, tissue type and abundance. Users can go through the three-dimensional presentations: rotate, pan and zoom to observe the data; move up or down through layers of information; and change filter settings at any point to expand or narrow the data set. According to Incyte, this visual approach not only makes it easier to investigate the gene sequences in the LifeSeq database, but also allows scientists to spot trends in gene expression and locate novel genes more easily. Incyte has used the program to identify gene fragments in LifeSeq that should be considered for inclusion in the LifeSeq FL database of full-length genes and the LifeSeq Atlas gene mapping database. In



Incyte Pharmaceuticals' LifeTools 3D software.

collaboration with Silicon Graphics, the company has created customized functions that are designed to help explore LifeSeq, including Protein Function Hierarchy Visualization, which displays the relative expression of genes across tissues. There is also MultiGene northern visualization, which uses a three-dimensional graph to show the tissue distribution for a selected set of genes. This program is a UNIX-based, client/server application, requiring Mineset 1.1 (offered by the company with LifeTools 3D).

Reader Enquiry No. 101

MM3 Software

From Tripos

The 1996 version of the company's molecular mechanics software

As with previous versions, this software can be accessed directly from the company's Sybyl suite of molecular design, visualization and analysis software. The company has also announced plans to expand the integration of its Sybyl suite with MM3. The plan is for MM3 to function as an energy server for Sybyl and other Tripos software packages, supplying molecular mechanics energies and force-field gradients on demand. In addition, Tripos is funding the development and testing of methods for MM3 to treat biomolecular structures, such as proteins, peptides and DNA.

Reader Enquiry No. 102

GeneWorld 2.1

Pangea Systems

A bioinformatics database application for automated sequence annotation and analysis

This product should find use in information-based pharmaceutical organizations seeking to use bioinformatics to discover new drug targets. GeneWorld 2.1 uses simplified, user-definable, automated workflow 'strategies', which enable automated analysis and anno-

tation of thousands of sequences. Once defined, these strategies can be saved and used for subsequent analyses, standardizing and distributing bioinformatics information throughout the user's organization. By selecting analysis steps from pull-down menus, customers can define strategies for automated gene finding, domain finding, functional identification, EST analysis and secondary structure prediction. A technology agreement with Paracel enables customers to include accelerated Smith-Waterman, pairwise alignment analyses in their strategies

Reader Enquiry No. 103

Gene Explorer

From MSI

A new computational environment for the analysis of DNA and protein sequences, as well as protein structures

From restriction analysis to virtual site-directed mutagenesis, Gene Explorer strives to meet the computational needs of molecular biology researchers within a single environment. It should help to focus the user's research on the most promising leads, hypotheses and candidate genes. In a series of steps, 3D Structure Prediction will transform protein sequence into a three-dimensional structure model, search databases for similar solved protein structures, present search results in a graphical display of homology, automatically align sequence homologues and highlight regions of similarity. Furthermore, it invokes Modeler algorithms to generate a model three-dimensional structure, verifying model structure using Profiles-3D technology. The process is presented in a simple, step-by-step interface, which is stated to enable users to build a model in a matter of minutes.

Reader Enquiry No. 104

Jaguar

From Schrödinger

Version 3.0 of the PS-GVB ab initio package has had its name changed to Jaguar

Jaguar includes a number of performance enhancements, with large speed-ups for density functional theory, as well as a new theoretical method, GVB-LMP2, which should give highly accurate conformational energies of molecules. Numerical methods are said to make this program a fast general-purpose program, which allows it to handle molecular systems previously considered intractable. According to the manufacturer, this package is useful for such operations as viewing electrostatic potentials of the different bind-

ing pockets of related proteins. The initial release included generalized valence bond (GVB) theory, which incorporated the pseudospectral (PS) numerical methods for fast evaluation of the two-electron integrals. Over the last three years, however, the product has outgrown its name in terms of functionalities and capabilities.

Reader Enquiry No. 105

Hardware systems

DeCypher II

From Time Logic

A system designed to provide the infrastructure to handle large-scale DNA sequence comparison analysis

The ultra-high performance cluster of this sequence similarity search system should help to identify rapidly similar sequences and deliver results quickly to the geneticist. The system's logic circuit adapts to additional algorithms through software reconfiguration, eliminating hardware obsolescence associated with older technologies. System specifications include 1.5–12 billion affine SW cells per second per rack, 1–8 15-accelerator systems per 48-cm rack, Xilinx reconfigurable computational arrays, 1–8 single- or dual-SMP Pentium CPUs, a Windows-NT TCP/IP network server, cluster scalability and a 32–144 GB hard-drive capacity per rack. CPU clusters offer high Smith–Waterman throughput, starting with a single 15-accelerator system. Each rack expands to hold up to eight systems, for a total of 120 Similarity Engine accelerators at a stated 12 billion affine SW cells per second. Special fault-tolerant versions are also available. The company offers custom hardware and software development and support. According to the manufacturer, virtually unlimited scalability is achieved by partitioning the query stream across one or more racks. Each 15-card CPU contributes up to 1,500

million affine SW cells per second. Users can reorganize the cluster for rapid interactive results, while production searches are streamed as a batch-mode, background process as resources become available. A simple, automatic script-based, semaphore mechanism pauses each CPU, permitting periodic database updates to be cloned across the cluster without loss of queued search requests or results. By using standard Intel technology with Windows-NT, the cost of ownership is lowered.

Reader Enquiry No. 106

Octane workstations

From Silicon Graphics

A new line that features high-performance graphics, symmetric multiprocessing and a pure 64-bit computing environment

Octane workstations are said to have improved systems and I/O bandwidth, giving greater flexibility to expand their applications. The system shares many components with the company's supercomputing architecture, employing a unique seven-port crossbar switch in lieu of the conventional shared bus architecture. It also combines this high-bandwidth switch with single or dual MIPS R10000 processors and fast desktop graphics. Furthermore, it is said to offer a 70 per cent increase in floating-point performance over the Indigo2 Impact 10000 family. With a peak bandwidth of 1.6 GB/s per port, the seven-port crossbar switch is stated to provide as much as 210 times the bandwidth found in traditional architectures. The system is designed to process multiple data types, including audio, video, imagery and three-dimensional geometry. With hardware-accelerated geometry processing and high-speed texturing and image processing, Octane systems enable real-time visual imagery for applications ranging from solid modelling, defence imaging, three-dimensional animation and emerging solu-

tions, such as real-time, three-dimensional volume roaming.

Reader Enquiry No. 107

More software

Synergy

From NetGenics

An alliance with Incyte Pharmaceuticals will link Synergy and Incyte's genomic database software

Synergy assists in the drug discovery process by enhancing the researcher's ability to access and analyse biological information. Incyte will incorporate certain aspects of NetGenics' software and the CORBA (common object request broker architecture) framework into its database products and the companies will jointly develop a product that is designed to provide increased functionality to Incyte's database and software subscribers by facilitating project management and allowing the integration and query of multiple, unrelated relational databases. Synergy is a cross-platform, client/server-based bioinformatics computer software system. It parallels the organization of multidisciplinary drug teams and enables the members of a drug team to communicate efficiently.

Reader Enquiry No. 108

Diva and RS³ Discovery

From Oxford Molecular

Six products designed to increase the speed and efficiency of the drug discovery process

These software packages add some 25 software tools for use in academic research and the pharmaceutical, chemical and biotechnology industries. Each product — Diva, CAChe 3.0, Tsar 3.0, Topkat 5.0, UniChem 4.0 and RS³ Discovery — deals with a different aspect of the process of drug discovery; taken together, they should help scientists predict the properties and toxicities of drug candidates and enable comparisons with

ADVERTISEMENTS

2-Hybrid System Media

Over 400 recipes and formulae in stock

CONVENIENCE

CONSISTENCY

ECONOMY

RELIABILITY

SELECTION

BIO 101 has the largest selection of Bacterial & Yeast Media available in capsule, powder & pouch form

Call us or your local distributor for ordering & information
(800) 424-6101 ♦ fax: (760) 598-0116
e-mail: info@bio101.com ♦ www.bio101.com

Gramsch Laboratories

New: Endomorphin -1 & -2 Custom Peptide Synthesis:
Shipped within one week:
10 AA DM 400.00; 15 AA DM 520.00;
20 AA DM 660.00; 10-30 mg unpurified;
buy only when satisfied with the HPLC profile

Antiserum Production Service:
you prepare the recombinant protein (in paa-gel or soluble), we produce the antiserum

Distributor for Japan: FUNAKOSHI CO, Tokio 113

Gramsch Laboratories
D-85247 SCHWABHAUSEN, Germany
Phone +49-8138-1769 or +49-8131-5679-0
Fax +49-8138-1260 or +49-8131-5679-24

CUSTOM HYBRIDOMA DEVELOPMENT

- Ascites Production
- Purification • Conjugation
- Polyclonal Antibodies

Also Available

1-800-481-9737

e-mail: antibodies@htbio.com
http://www.htbio.com
fax: (760) 788-9694
P.O. Box 1319
Ramona, CA 92065

HTI
HTI Bio-Products, Inc.

chemicals in existing databases. Diva was created to allow researchers to mine corporate databases and identify the properties and relationships of thousands of compounds. Version 3.0 of CAChe allows experimental chemists to predict and analyse properties of drug candidates and chemicals on a PC. Tsar 3.0 was written to help identify structure-activity relationships quickly by integrating QSAR data, computations and statistics into one package. Topkat, now in version 5.0, offers a Windows environment and focuses experimental testing by quickly and accurately assessing the toxicity of drug candidates based on structure. UniChem optimizes chemical reactions with property predictions that should aid in the design of catalysts and in the prediction of reaction mechanisms. Lastly, RS³ Discovery uses Oracle to store and retrieve information in a flexible open architecture.

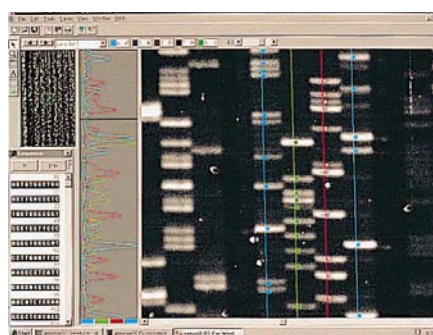
Reader Enquiry No. 109

extendLRS

From Genomyx

Software for automated DNA sequencing and analysis

The extendLRS sequencing software has been designed primarily for the analysis of



extendLRS sequencing software from Genomyx.

images obtained from the Genomyx SC fluorescent imaging system. It also processes images from other sources, including autorads and images as large as 50 Mb. This system includes the company's automatic band detecting and labelling, and lane tracking technology.

By eliminating manual calculations and guesswork, extendLRS increases base-calling accuracy, says Genomyx. The new software can read sequences as long as 400 bases with a high degree of accuracy. Several software tools are included for stepping through and editing the sequence. DNA sequence information can be exported in ASCII text or SCF format for use in other

applications — TIFF images may be imported or exported. Images and sequences can be printed on any standard printer as they are or with lanes, band labels and annotation.

Reader Enquiry No. 110

Antisense Functional Genomics

From Sequitur

The commercial availability of second-generation antisense compounds

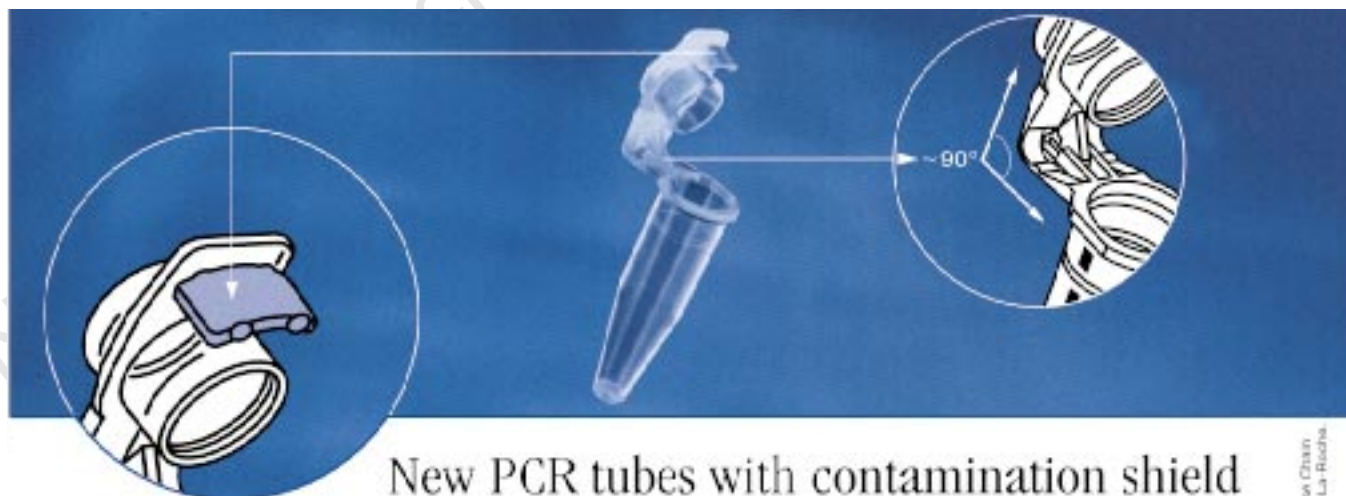
These compounds have been developed through careful screening in a cell culture assay system. With this program, the company does not retain the rights to the research findings derived through their use.

Advantages of these compounds are said to include improved specificity and reduced toxicity, target gene inhibition with only three to six compounds screened, as well as rapid and cost-effective results, according to Sequitur.

Reader Enquiry No. 111

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENT



● Accessories: The work tray and the accompanying frame may be used as a tube rack or for inserting tubes into the thermoblock.



The innovative contamination shield on the lid of the new 0.2 ml PCR tube prevents accidental contact with the inner surface of the lid during opening and closing.

The tubes also have a special hinge which allows the lid to click into a set position of approx. 90° and thus prevents overlapping with adjacent tubes.

The new, thin-walled PCR tubes from Eppendorf are available as individual tubes, and also as strips of five or eight tubes.

New Eppendorf Thermocyclers: Mastercycler[®] gradient and Mastercycler[®] personal. Ask for your information pack now!

eppendorf

Eppendorf-Netheler-Hinz GmbH · 22331 Hamburg · Germany · Tel. +49 (40) 5 38 01-0 · Fax +49 (40) 5 38 01-595
e-mail: eppendorf@eppendorf.com · eppendorf home page: <http://www.eppendorf.com>

eppendorf[®] is a registered trademark. PCR (Polymerase Chain Reaction) is protected by patent owned by Hoffmann-La Roche.